

Thu Apr 24 08:54:04 2003

PT New yeast inhibitory peptide useful for inhibiting viral protein
PT translation and replication -
XX
PS Claim 5; Page 57; 81pp; English.

XX This sequence is the La autoantigen binding domain (LAP). LAP is a
CC cellular protein which binds to an internal ribosome entry site (IRES).
CC The peptide is used to inhibit mRNA translation, viral protein
CC translation or viral replication. Viral mRNA translation is initiated at
CC the IRES and the LAP peptide prevents translation initiation factors
CC from binding at the site. The peptide can be used alone or in combination
CC with an inhibitor RNA (RNA see A4245200). The LAP peptide is useful as
CC an antiviral agent, which works through the inhibition of mRNA
CC translation, especially viral mRNA. Examples of viruses which can be
CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
CC viruses, rabdovirus, adenovirus, and parainfluenza virus, poliovirus,
CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
CC The peptide also inhibits replication of the above viruses. The LAP
CC peptide selectively inhibits viral protein translation, and is therefore
CC not toxic to the host cell.

XX SQ Sequence 18 AA;
Query Match 90.6%; Score 87; DB 21; Length 18;
Best Local Similarity 88.9%; Pred. No. 1.6e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQICQIEYYFGDF 18
DB 1 AALEAKICHQIEYYFGDF 18

RESULT 2
AAR43394
ID AAR43394 standard; peptide; 21 AA.

XX AC AAR43394;
XX 12-MAY-1994 (first entry)
XX La/SSB epitope 17.
XX Linear; epitope: 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

XX OS Homo sapiens.
XX WO9321223-A.
XX 28-OCT-1993.
XX 13-APR-1993; 93WO-US03484.
XX 13-APR-1992; 92US-0867819.
XX (OKLA) UNIV OKLAHOMA STATE.
XX Harley JB;

XX WPI; 1993-351658/44.
XX New linear epitope(s) for human auto-antibodies - from the
PT Ro/SSA, La/SSB and Sm B/B' antigens and ribo:nucleoprotein, used
PT for diagnosing and treating auto-immune disorders e.g. systemic
PT lupus-erythematosus

XX Claim 1; Page 30; 43pp; English.
XX The sequences given in AAR43391-562 are linear epitopes which are

CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
CC polypeptide. These antigens are common in systemic lupus
CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
CC of proteins has been shown to have several molecular forms which are
CC defined by the molecular weight of the antigen identified. The major
CC form has a molecular weight of 60 kD and two additional forms have
CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
CC group of autoantibodies and binds small RNAs with a polypyridine
CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
CC phosphoprotein which associates with RNA polymerase III transcripts.
CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
CC U5 RNA. Anti-Sm antibodies may be directed against one or a
CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
CC used for preventing, treating or screening autoimmune disorders,
CC especially SLE or Sjogrens syndrome (SS). They bind to a human
CC autoantibody and may therefore be used as vaccines.

XX SQ Sequence 21 AA;

Query Match 67.7%; Score 65; DB 14; Length 21;
Best Local Similarity 91.7%; Pred. No. 0.00077;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICQIEYYFGDF 18
DB 1 ICHQIEYYFGDF 12

RESULT 3
AAU00924
ID AAU00924 standard; Peptide; 23 AA.

XX AC AAU00924;
XX 18-MAY-2001 (first entry)
XX VH ligand-binding domain variant CDR3 region #31.

XX Human; monoclonal antibody A6; VH ligand-binding domain; variable region;
KW parental ligand binding molecule; PLBW; framework region; FR; CDR;
KW combinatorial library; diagnosis; therapy; surface protein; pesticide;
KW cell-receptor; carbohydrate; lipid; antibiotic; steroid; hormone; dye;
KW herbicide; polymer; immunoglobulin; CDR3 region; M2 immunoglobulin;
KW panning.

XX OS Homo sapiens.
XX WO200118058-A2.
XX 15-MAR-2001.
XX 07-SEP-2000; 2000WO-CA01027.
XX 07-SEP-1999; 99CA-2282179.
XX 04-NOV-1999; 99US-0163546.
XX (NOVO-) NOVOPHARM BIOTECH INC.

XX Kaplan H, Entwistle J, Tanha J, Narang S, Dan M;
XX WPI; 2001-235191/24.

XX Combinatorial libraries including phage display library comprises
PT variants of immunoglobulin VH fragments which comprises the framework
PT regions of wild-type or modified immunoglobulin VH domain of human A6
PT antibody

XX Disclosure; Page 25; 133pp; English.

XX The sequence represents a variant CDR3 region of a parental VH

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:28:18 ; Search time 27.2022 Seconds
(without alignments)
88.173 Million cell updates/sec

Title: US-09-836-073-2

Perfect score: 96
Sequence: 1 AALEAICQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 320064

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

ID	Score	Query Match	Length	ID	Description
1:	87	90.6	18	21	AAV52200
2:	65	67.7	23	22	AAAR43394
3:	34	35.4	21	22	AAU00924
4:	33	34.4	18	20	AAW93362
5:	32	33.3	14	22	AAW97227
6:	32	33.3	20	18	AAW42900
7:	32	33.3	21	22	AAW69417
8:	32	33.3	22	20	AAV19621
9:	32	33.3	25	17	AAW92123
10:	32	33.3	25	20	AAW80367
11:	87	90.6	18	21	AAV52200
12:	65	67.7	23	22	AAAR43394
13:	34	35.4	21	22	AAU00924
14:	33	34.4	18	20	AAW93362
15:	32	33.3	14	22	AAW97227
16:	32	33.3	20	18	AAW42900
17:	32	33.3	21	22	AAW69417
18:	32	33.3	22	20	AAV19621
19:	32	33.3	25	17	AAW92123
20:	32	33.3	25	20	AAW80367
21:	87	90.6	18	21	AAV52200
22:	65	67.7	23	22	AAAR43394
23:	34	35.4	21	22	AAU00924
24:	33	34.4	18	20	AAW93362
25:	32	33.3	14	22	AAW97227
26:	32	33.3	20	18	AAW42900
27:	32	33.3	21	22	AAW69417
28:	32	33.3	22	20	AAV19621
29:	32	33.3	25	17	AAW92123
30:	32	33.3	25	20	AAW80367

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	90.6	18	21	AAV52200
2	65	67.7	23	22	AAAR43394
3	34	35.4	21	22	AAU00924
4	33	34.4	18	20	AAW93362
5	32	33.3	14	22	AAW97227
6	32	33.3	20	18	AAW42900
7	32	33.3	21	22	AAW69417
8	32	33.3	22	20	AAV19621
9	32	33.3	25	17	AAW92123
10	32	33.3	25	20	AAW80367

11	31	32.3	8	14	AAAR43395	La/SSB epitope 24.
12	31	32.3	10	23	AAE22202	Murine MC-1 antio
13	31	32.3	19	10	AAAP90453	Epitope recognised
14	31	32.3	21	23	ABG66356	IgE Fcεpsilon RI b
15	31	32.3	23	22	AAU00592	VH ligand-binding
16	30	31.2	11	15	AAAS6292	Synthetic derivati
17	30	31.2	14	22	AAAG6456	Human olfactory re
18	30	31.2	21	22	ABH27798	Human peptide #449
19	30	31.2	21	22	ABH32969	Peptide #475 encod
20	30	31.2	21	22	ABH18443	Protein #442 encod
21	30	31.2	21	22	AAH53767	Human brain expres
22	30	31.2	21	22	AAH66153	Human bone marrow
23	30	31.2	21	22	AAH14024	Peptide #458 encod
24	30	31.2	21	22	AAH26430	Peptide #467 encod
25	30	31.2	21	22	AAH01765	Peptide #447 encod
26	30	31.2	21	23	ABG35803	Human peptide enco
27	30	31.2	24	14	AAH36918	Insulin-like growt
28	29	30.2	9	21	AAH82311	Antibody binding H
29	29	30.2	10	15	AAH56297	Synthetic modified
30	29	30.2	11	15	AAH56307	Modified Influenza
31	29	30.2	11	20	AAH56255	Anti-progesterone
32	29	30.2	13	22	ABH52265	Human API-75 trypt
33	29	30.2	13	22	AAU28698	DPI tryptic digest
34	29	30.2	13	22	AAU28716	DPI tryptic digest
35	29	30.2	13	22	AAU25223	Schizophrenia-Asso
36	29	30.2	13	22	AAU26364	Depression-Asso
37	29	30.2	13	22	AAU26364	Depression-Asso
38	29	30.2	13	22	AAU15567	Schizophrenia-asso
39	29	30.2	13	22	AAH87228	Breast-cancer asso
40	29	30.2	14	17	AAH90665	S. acidocaldarius
41	29	30.2	18	22	AAH20426	Anti-FIX/FIXa anti
42	29	30.2	19	21	AAH59094	Breast and ovarian
43	29	30.2	21	23	ABG66499	IgE Fcεpsilon RI b
44	29	30.2	25	22	AAH66069	Human TANGO 294 tr
45	28	29.2	9	23	ABP47453	N. meningitidis LO

ALIGNMENTS

RESULT 1

AAV52200
ID AAV52200 standard; peptide; 18 AA.

XX AAV52200;
AC AAV52200;

XX DT 14-MAR-2000 (first entry)
XX DE Human la autoantigen peptide (LAP).

XX DE Human la autoantigen peptide (LAP).
KW La autoantigen; LAP; internal ribosome entry site; IRES; translation;

KW viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
KW coronavirus; hepatitis virus; rhinovirus; adenovirus; coxsackie virus;
KW parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
KW vesicular stomatitis virus.
XX Homo sapiens.

XX OS Homo sapiens.
PN WO9961613-A2.

XX PD 02-DEC-1999.
XX PF 21-MAY-1999; 99WO-US11281.

XX PR 22-MAY-1998; 98US-0086527.
XX PA (REGC) UNIV CALIFORNIA.

XX PI Das S, Dasgupta A;
XX DR WPI; 2000-062712/05.

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:42 ; Search time 4.95506 Seconds
(without alignments)
150.669 Million cell updates/sec

Title: US-09-836-073-2
Perfect score: 96
Sequence: 1 AALEAQICQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	90.6	404	LA_BOVIN	P10881 bos taurus
2	87	90.6	408	LA_HUMAN	P05455 homo sapien
3	87	90.6	415	LA_RAT	P38656 rattus norv
4	83	86.5	415	LA_MOUSE	P32067 mus musculu
5	68	70.8	427	LAB_XENLA	P28049 xenopus lae
6	64	66.7	428	LAA_XENLA	P28048 xenopus lae
7	56	58.3	383	LA_AEDAL	Q26457 aedes albop
8	52	54.2	390	LA_DROME	P40796 drosophila
9	49	51.0	298	LAHL_SCHPO	P87058 schizosacch
10	47	49.0	506	TCMO_PETCR	Q43033 petroselinu
11	44	45.8	392	GALI_HUMAN	P51570 homo sapien
12	41	42.7	466	SRO9_YEAST	P25567 saccharomyc
13	40	41.7	176	IPYR_HAEIN	P44529 haemophilus
14	40	41.7	292	PPNK_ECO57	P58057 escherichia
15	40	41.7	292	PPNK_ECOLI	P37768 escherichia
16	40	41.7	391	GALI_MOUSE	Q90n00 mus musculu
17	40	41.7	505	TCMO_ARATH	P92994 arabidopsis
18	40	41.7	505	TCMO_CICAR	Q81928 cicor ariet
19	40	41.7	505	TCMO_GLYEC	Q96423 glycyrrhiza
20	40	41.7	505	TCMO_HELTU	Q04468 helianthus
21	40	41.7	505	TCMO_PEA	Q43067 pisum sativ
22	40	41.7	505	TCMO_PHAU	P37115 phaseolus a
23	40	41.7	505	TCMO_ZINEL	Q43240 zinnia eleg
24	40	41.7	506	TCMO_MEDSA	P37114 medicago sa
25	40	41.7	506	TCMO_SOYBN	Q42797 glycine max
26	40	41.7	646	YE14_SCHPO	Q13869 schizosacch
27	40	41.7	1180	C4AA_BACTI	P16480 bacillus th
28	40	41.7	1418	CELL_CAEEL	P34641 caenorhabdi
29	39	40.6	363	LE3A_ASPNG	P87256 aspergillus
30	39	40.6	475	VG41_BPT4	P04530 bacterioph
31	39	40.6	482	LBP_RASIT	P17454 oryctolagus
32	39	40.6	490	IFT4_HUMAN	Q14879 homo sapien
33	39	40.6	1164	RPO2_COMFX	P17474 cowpox viru

34	39	40.6	1164	1	RPO2_VACCV	P19798 vaccinia vi
35	39	40.6	1164	1	RPO2_VARV	P33811 variola vir
36	39	40.6	1176	1	NIR_NEUCR	P38681 neurospora
37	38.5	40.1	410	1	NEUS_CHICK	Q90935 gallus gall
38	38.5	40.1	410	1	NEUS_HUMAN	Q99574 homo sapien
39	38.5	40.1	410	1	NEUS_MOUSE	O35684 mus musculu
40	38.5	40.1	410	1	NEUS_RAT	Q911d2 rattus norv
41	38.5	40.1	503	1	GLPK_PSETO	O87924 pseudomonas
42	38	39.6	177	1	IPYR_NEIMA	Q91jv3 neisseria m
43	38	39.6	192	1	RR42_CVACA	O22029 cyanidium c
44	38	39.6	275	1	LAH1_YEAST	P33399 saccharomyc
45	38	39.6	450	1	G28A_DROME	Q9vm09 drosophila

ALIGNMENTS

RESULT 1				
LA_BOVIN				
ID	LA_BOVIN	STANDARD;	PRT;	404 AA.
AC	P10881;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-JUL-1989 (Rel. 11, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Lupus Ia protein homolog (La ribonucleoprotein) (La autoantigen homolog).			
GN	SSB.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=pituitary;			
RX	MEDLINE=89202037; PubMed=2468131;			
RA	Chan E.K.L., Sullivan K.F., Tan E.M.;			
RT	"Ribonucleoprotein SS-B/La" belongs to a protein family with consensus sequences for RNA-binding.;			
RL	Nucleic Acids Res. 17:2233-2244(1989).			
CC	-!- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S, AND 7-2 RNAS.			
CC	-!- SUBCELLULAR LOCATION: Nuclear (Probable).			
CC	-!- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE C-TERMINAL PART OF THE PROTEIN.			
CC	-!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).			

Query Match 90.6% Score 87; DB 1; Length 404;

Best Local Similarity 88.9%; Pred. No. 1.3e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 AALEAQICQIEYFGDF 18
Db 11 AALEAKICHQIEYFGDF 28

RESULT 2

LA_HUMAN STANDARD; PRT; 408 AA.
AC P05455;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lupus La protein (Stogren syndrome type B antigen (SS-B)) (La
DE ribonucleoprotein) (La autoantigen).
GN SSB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89202037; PubMed=2468131;
RA Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
RT sequences for RNA-binding.";
RL Nucleic Acids Res. 17:2233-2244(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89053970; PubMed=3192525;
RA Chambers J.C., Renan D., Martin S.J., Keene J.D.;
RT "Genomic structure and amino acid sequence domains of the human La
RT autoantigen.";
RL J. Biol. Chem. 263:18043-18051(1988).
RN [3]
RP SEQUENCE OF 54-408 FROM N.A.
RX MEDLINE=88199081; PubMed=2452201;
RA Sturgess A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
RA Coppel R.S.;
RT "Characteristics and epitope mapping of a cloned human autoantigen
RT La.";
RL J. Immunol. 140:3212-3218(1988).
RN [4]
RP SEQUENCE OF 54-97 FROM N.A.
RX MEDLINE=85166283; PubMed=3856888;
RA Chambers J.C., Keene J.D.;
RT "Isolation and analysis of cDNA clones expressing human lupus La
RT antigen.";
RL proc. Natl. Acad. Sci. U.S.A. 82:2115-2119(1985).
RN [5]
RP FUNCTION.
RX MEDLINE=89251617; PubMed=2470590;
RA Gottlieb E., Steltz J.A.;
RT "Function of the mammalian La protein: evidence for its action in
RT transcription termination by RNA polymerase III.";
RL EMBO J. 8:851-861(1989).
RN [6]
RP PHOSPHORYLATION.
RX MEDLINE=97207017; PubMed=9054510;
RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Marale R.J.;
RT "Phosphorylation of the human La antigen on serine 366 can regulate
RT recycling of RNA polymerase III transcription complexes.";
RL Cell 88:707-715(1997).
RN [7]
RP -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
CC AND 7-2 RNAs.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE

C-TERMINAL PART OF THE PROTEIN.
-1- PTM: THE N-TERMINUS IS BLOCKED.
-1- DISEASE: SERA FROM PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS
CC OFTEN CONTAIN ANTIBODIES THAT REACT WITH THE NORMAL CELLULAR
CC LA PROTEIN AS IF THESE ANTIGEN WAS FOREIGN.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X13697; CAA31985.1; -
CC EMBL; J04205; AAA51885.1; -
DR PIR; A31888; A31888.
DR PIR; A22956; A22956.
DR PIR; A31273; A31273.
DR PIR; S03848; S03848.
DR PIR; S11013; S11013.
DR Genew; HGNC:11316; SSB.
DR MIN; 109090; -
DR InterPro; IPR002344; Lupus.La.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KW Systemic lupus erythematosus; RNA-binding; Phosphorylation;
KW Nuclear protein.
FT DOMAIN 111 187 RNA-BINDING (RRM).
FT MOD_RES 366 366 PHOSPHORYLATION (BY CK2).
SQ SEQUENCE 408 AA; 46837 MW; ECL53C15F9187FC4 CRC64;
Query Match 90.6%; Score 87; DB 1; Length 408;
Best Local Similarity 88.9%; Pred. No. 1.3e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 AALEAQICQIEYFGDF 18
Db 11 AALEAKICHQIEYFGDF 28

RESULT 3

LA_RAT STANDARD; PRT; 415 AA.
AC P36556;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
DE homolog).
GN SSB OR SS-B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93245255; PubMed=7916708;
RA Sensel I., Troester H., Bartsch H., Schwemmler M., Igloi G.L.,
RA Bachmann M.;
RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La:
RT detection of species-specific variations.";
RL Gene 126:265-268(1993).
CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,

```
CC AND 7-2 RNAS.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X67859; CAA48043.1; -
DR PIR; JCI494;
DR InterPro; IPR002344; Lupus_La
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KW RNA-binding; Nuclear protein; Phosphorylation.
FT DOMAIN 111 187 RNA-BINDING (RRM).
SQ SEQUENCE 415 AA; 47777 MW; 033FD9CC1E475F98 CRC64;

Query Match 90.6%; Score 87; DB 1; Length 415;
Best Local Similarity 88.9%; Pred. No. 1.3e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQICQIQEYFGDF 18
DB 11 AALEAKICHQIEYFGDF 28

RESULT 4
LA_MOUSE STANDARD; PRT; 415 AA.
AC P32067;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
DE homolog).
GN SSB OR SS-B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93203630; PubMed=8454877;
RA Topfer F., Gordon T., McCluskey J.;
RT "Characterization of the mouse autoantigen La (SS-B). Identification
RT of conserved RNA-binding motifs, a putative ATP binding site and
RT reactivity of recombinant protein with poly(U) and human
RT autoantibodies."
RL J. Immunol. 150:3091-3100(1993).
RN [2]
RP SEQUENCE OF 1-11 FROM N.A.
RA Groelz D., Bachmann M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
CC AND 7-2 RNAS.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L00993; AAA39415.1; -
DR EMBL; Y07951; CAA69249.1; -
DR MGD; MGI:98423; Ssb
DR InterPro; IPR002344; Lupus_La
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KW RNA-binding; Nuclear protein; Phosphorylation.
FT DOMAIN 111 187 RNA-BINDING (RRM).
SQ SEQUENCE 415 AA; 47756 MW; 2D75197692FDC933 CRC64;

Query Match 86.5%; Score 83; DB 1; Length 415;
Best Local Similarity 88.2%; Pred. No. 6.5e-07;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALEAQICQIQEYFGDF 18
DB 12 ALEAKICHQIEYFGDF 28

RESULT 5
LAB_XENLA STANDARD; PRT; 427 AA.
AC P28049;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen
DE homolog B).
GN LABI.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Oocyte;
RX MEDLINE=93287095; PubMed=8510143;
RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
RT "La proteins from Xenopus laevis. cDNA cloning and developmental
RT expression."
RL J. Mol. Biol. 231:196-204(1993).
CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- DEVELOPMENTAL STAGE: BARELY DETECTABLE IN STAGE I/II OOCYTES,
CC ACCUMULATE IN STAGE III/IV OOCYTES, THEN EXHIBIT A ROUGHLY
CC CONSTANT STEADY STATE LEVEL IN MATURE OOCYTES, EGGS, AND EARLY
CC EMBRYOS.
CC -1- PTM: PHOSPHORYLATED (PROBABLE).
CC -1- MISCELLANEOUS: THERE ARE TWO FORMS OF LA, LAA AND LAB, IN XENOPUS.
CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```


Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zhong X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
The genome sequence of *Drosophila melanogaster*.;
Nature 387:2185-2195(2000).
CC CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTION TERMINATION BY RNA
POLYMERASE III. BINDS RNA AND DNA. BINDS TO PRECURSORS OF RNA POLYMERASE III TRANSCRIPTS. MAY PLAY A SPECIALIZED ROLE DURING FLX DEVELOPMENT.
CC CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT EMBRYONIC, LARVAL, PUPAL, AND ADULT DEVELOPMENT. EXPRESSION THROUGHOUT THE EMBryo IS FOLLOWED BY A RESTRICTED PATERN OF MESODERMAL EXPRESSION THAT IS LATER CONFINED TO THE VISERAL MESODERM, GNADS, GUT, AND SALIVARY GLANDS.
CC CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See http://www.isb-sib.ch/announce or send an email to licensee@isb-sib.ch).

EMBL; U07652; AAA20518.1; --
DR EMBL; L32988; AAA21776.1; --
DR EMBL; AE003666; AA053885.1; --
DR FlyBase; FBgn0011638; La.
DR InterPro; IPR002344; Lupus_La.
DR IntePro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm_1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KW RNA-binding; Nuclear protein; DNA-binding.
FT DOMAIN 149 234 RNA-BINDING (RRM)..
FT CONFLICT 169 169 A->T (IN REF. 1).
FT CONFLICT 182 183 KH->NS (IN REF. 1).
FT CONFLICT 283 283 A->R (IN REF. 1).
FT CONFLICT 329 329 K->N (IN REF. 1).
SQ SEQUENCE 390 AA; 44884 MW; A8099288B90446A5 CRC64;

Query Match 54.28; Score 52; DB 1; Length 390;
Best local Similarity 64.3%; Pred.No. 0.16;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 EAQTQQIETYYFGD 17
ID ID::IIIII
Db 51 ERAIRQEYYFGD 64

RESULT 9
LAHL_SCHPO STANDARD; PRT; 298 AA.
AC P87058; Q10458; O13362;
DT 15-JUL-1998 (Rel. 36, Created)
DI 15-JUL-1998 (Rel. 36, Last sequence update)
DD 15-JUN-2002 (Rel. 41, Last annotation update)
DE The protein homolog (La ribonucleoprotein) (La autoantigen homolog).
GS SLAI OR SPACS7A10.10C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_Taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98067398; PubMed=9404894;

RA van Horn D.J., Yoo C.J., Xue D., Shi H., Wolin S.L.;
 RT "The La protein in Schizosaccharomyces pombe: a conserved yet
 RT dispensable phosphoprotein that functions in tRNA maturation";
 RL RNA 3:1434-1443(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Utsumi R.R.U.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Raydread M.A., Lyne M., Lyne R., Stewart A.,
 RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman C.M.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feldwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Widdjens I., Volckaert G., Wertz R., Robben J., Grymonprez B.,
 RA Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: BINDS TO THE PRECURSORS OF POLYMERASE III RNAS.
 CC -1- FUNCTIONS IN TRNA MATURATION.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF022949; AB82145.1; -
 CC DR EMBL: AB011371; BAA24981.1; -
 CC DR EMBL: Z94864; CAB08173.1; -
 CC DR InterPro: IPR002344; Lupus_La.
 CC DR InterPro: IPR000504; RNA_rec_mot.
 CC DR Pfam: PF00076; rrm; 1.
 CC DR PRINTS: PR00302; LUPUSLA.
 CC DR SMART: SM00360; RRM; 1.
 CC DR PROSITE: PS0102; RRM; 1.
 CC DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 KW RNA-binding; Nuclear protein.
 FT DOMAIN 154 236 RNA-BINDING (RRM).
 FT CONFLICT 188 188 M -> I (IN REF. 1 AND 2).
 SQ SEQUENCE 298 AA; 34616 MW; 64E6AB99940B87F4 CRC64;
 Query Match 51.0%; Score 49; DB 1; Length 298;
 Best Local Similarity 50.0%; Pred. No. 0.39;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 QY 4 EAQICQIQIEYFGD 17
 ||:::|:|:|

Db 64 EAEVLKQVEFFPSD 77
 RESULT 10
 TCMO_PETCR STANDARD; PRT; 506 AA.
 ID Q43033;
 AC 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trans-cinnamate 4-monooxygenase (EC 1.14.13.11) (Cinnamic acid
 DE 4-hydroxylase) (CAAH) (C4H) (P450C4H) (Cytochrome P450 73).
 GN CYP7A10 OR CYP73.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
 OX NCBI_TaxID=4043;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95320184; PubMed=7597051;
 RA Logemann E., Parniske M., Hahlbrock K.;
 RT "Modes of expression and common structural features of the complete
 RT phenylalanine ammonia-lyase gene family in parsley";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:5905-5909(1995).
 CC -1- FUNCTION: CONTROLS CARBON FLUX TO PIGMENTS ESSENTIAL FOR
 CC POLLINATION OR UV PROTECTION, TO NUMEROUS PYTOALEXINS SYNTHESIZED
 CC BY PLANTS WHEN CHALLENGED BY PATHOGENS, AND TO LIGNINS.
 CC -1- CATALYTIC ACTIVITY: Trans-cinnamate + NADPH + O(2) -> 4-
 CC hydroxycinnamate + NADP(+) + H(2)O.
 CC -1- PATHWAY: Phenylpropanoid metabolism; second step.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L38898; AAC41660.1; -
 CC DR InterPro: IPR001128; Cytochrome_P450.
 CC DR Pfam: PF00067; P450; 1.
 CC DR PRINTS: PR00385; P450.
 CC DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Heme; NADP.
 FT BINDING 448 448 HEME (BY SIMILARITY).
 SQ SEQUENCE 506 AA; 58047 MW; 32F00E959D69CCF CRC64;
 Query Match 49.0%; Score 47; DB 1; Length 506;
 Best Local Similarity 50.0%; Pred. No. 1.5;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 5 AQICQIQIEYFGDF 18
 ::::|:|:|:|
 Db 215 SRLAQSFYHFGDF 228
 RESULT 11
 GALL_HUMAN STANDARD; PRT; 392 AA.
 ID P51570;
 AC 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Galactokinase (EC 2.7.1.6) (Galactose kinase).
 GN GALK1 OR GALK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

AC P44529;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
 DE hydrolase) (PPase).
 GN PPA OR H10124.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512(1995).
 CC -|- CATALYTIC ACTIVITY: Diphosphate + H(2)O -> 2 phosphate.
 CC -|- COFACTOR: REQUIRES THE PRESENCE OF DIVALENT METAL CATION.
 CC MAGNESIUM CONFERS THE HIGHEST ACTIVITY. BINDS 4 DIVALENT CATIONS
 CC PER SUBUNIT (BY SIMILARITY).
 CC -|- SUBUNIT: HOMOHXAMER (BY SIMILARITY).
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -|- SIMILARITY: BELONGS TO THE PPASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32698; AAC21798.1;
 CC DR HSSP: P50308; IQEZ.
 CC DR TIGR: H10124;
 CC DR InterPro: IPR001596; Pyrophosphatase.
 CC DR Pfam: PF00719; Pyrophosphatase; 1.
 CC DR ProDom: PD002014; Pyrophosphatase; 1.
 CC DR PROSITE: PS00387; PPASE; 1.
 CC KW Hydrolase; Magnesium; Complete proteome.
 FT ACT_SITE 31 31 BY SIMILARITY.
 SQ SEQUENCE 176 AA; 19725 MW; ECDDDE557F28446F CRC64;
 Query Match 41.7%; Score 40; DB 1; Length 176;
 Best Local Similarity 38.9%; Pred. No. 8.2;
 Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
 QY 1 AALEAQICQOIYYFGDF 18
 DB 125 ADLPANLIKQIEPFNNY 142
 RESULT 14
 PPNNK_ECO57 STANDARD; PRT; 292 AA.
 ID PPNNK_ECO57
 AC P58057;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable inorganic polypophosphate/ATP-NAD kinase (EC 2.7.1.23)
 DE (Poly(P)/ATP NAD kinase).
 GN PPNNK OR Z3908 OR ECS3477.

OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocous K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 CC -|- FUNCTION: Catalyzes the phosphorylation of NAD to NADP. Utilizes
 CC ATP and other nucleoside triphosphates as well as inorganic
 CC polyphosphate as a source of phosphorus (By similarity).
 CC -|- CATALYTIC ACTIVITY: ATP + NAD(+) -> ADP + NADP(+).
 CC -|- COFACTOR: Requires divalent metal ions for activity (By
 CC similarity).
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -|- SIMILARITY: BELONGS TO THE NAD KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF005491; BAG57725.1;
 CC DR EMBL: AF002562; BAB36900.1;
 CC DR InterPro: IPR002504; ATP_NADK.
 CC DR Pfam: PF01513; NAD_kinase; 1.
 CC KW Transferase; Kinase; NAD; NADP; Complete proteome.
 SQ SEQUENCE 292 AA; 32596 MW; 30BBE44DA060915 CRC64;
 Query Match 41.7%; Score 40; DB 1; Length 292;
 Best Local Similarity 87.5%; Pred. No. 14;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 3 LEAQICQQ 10
 DB 129 LEAQVCQQ 136
 RESULT 15
 PPNNK_ECOLI STANDARD; PRT; 292 AA.
 ID PPNNK_ECOLI
 AC P37768; P46140; P77490;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable inorganic polypophosphate/ATP-NAD kinase (EC 2.7.1.23)
 DE (Poly(P)/ATP NAD kinase).
 GN PPNNK OR B2615.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.

OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of *Escherichia coli* K-12";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshima T., Oyama S., Saito N., Sempel G., Satoh Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the *Escherichia coli*
RT - K12 genome corresponding to 50.0-58.8 min on the linkage map and
RT analysis of its sequence features";
RL DNA Res. 4:91-113(1997).
RN [3]
RP SEQUENCE OF 1-163 FROM N.A.
RC STRAIN-B178;
RX MEDLINE=88319942; PubMed=3045760;
RA Lipinska B., King J., Ang D., Georgopoulos C.;
RT "Sequence analysis and transcriptional regulation of the *Escherichia*
RT coli *gpe* gene, encoding a heat shock protein.";
RL Nucleic Acids Res. 16:7545-7562(1988).
RN [4]
RP SEQUENCE OF 194-292 FROM N.A.
RC STRAIN-K12;
RX MEDLINE=87259985; PubMed=3037486;
RA Rostas K., Morton S.J., Pickles S.M., Lloyd R.G.;
RT "Nucleotide sequence and LexA regulation of the *Escherichia coli* *recN*
RT gene";
RL Nucleic Acids Res. 15:5041-5049(1987).
RN [5]
RP IDENTIFICATION.
RX MEDLINE=95075659; PubMed=7984428;
RA Borodovsky M., Rudd K.E., Koonin E.V.;
RT "Intrinsic and extrinsic approaches for detecting genes in a
RT bacterial genome";
RL Nucleic Acids Res. 22:4756-4767(1994).
CC -1- FUNCTION: Catalyzes the phosphorylation of NAD to NADP. Utilizes
CC ATP and other nucleoside triphosphates as well as inorganic
CC polyphosphate as a source of phosphorus (By similarity).
CC -1- CATALYTIC ACTIVITY: ATP + NAD(+) -> ADP + NADP(+).
CC -1- COFACTOR: Requires divalent metal ions for activity (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE NAD KINASE FAMILY.
CC -1- CAUTION: REF.3 AND REF.4 SEQUENCES DIFFER FROM THAT SHOWN DUE TO
CC FRAMESHIFTS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC EMBL; AE000347; AAC75664.1; -;
DR EMBL; D90888; BAA16500.1; -;
DR EMBL; X07863; -; NOT_ANNOTATED_CDS.
DR EMBL; U36840; AAA9785.1; -;
DR EMBL; Y00357; -; NOT_ANNOTATED_CDS.
DR EcoGene; EG12192; ppnk.

DR InterPro; IPR002504; ATP_NADK.
DR Pfam; PF01513; NAD_Kinase; 1.
KW Transferase; Kinase; NAD; NADP; Complete proteome.
SQ SEQUENCE 292 AA; 32566 MW; D1E631658408F2E1 CRC64;
Query Match 41.7%; Score 40; DB 1; Length 292;
Best Local Similarity 87.5%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 LEAQICQQ 10
Db 129 LEAQVCCQ 136
Search completed: April 23, 2003, 13:28:03
Job time : 5.95506 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:03 ; Search time 10.1124 Seconds
(without alignments)
171.119 Million cell updates/sec

Title: US-09-836-073-2
Perfect score: 96
Sequence: 1 ALEAQICQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	90.6	404	1 S03849	ribonucleoprotein
2	87	90.6	408	1 A31888	ribonucleoprotein
3	87	90.6	415	1 JC1494	ribonucleoprotein
4	68	70.8	427	1 S33817	ribonucleoprotein
5	64	66.7	428	1 S33818	ribonucleoprotein
6	52	54.2	390	2 A53773	La/SS-B homolog D-
7	52	54.2	390	2 A53781	ribonucleoprotein
8	50	52.1	391	2 H89777	capsular polysacch
9	49	51.0	298	2 T38937	rna binding protei
10	49	51.0	298	2 T43542	RNA-binding protei
11	47	49.0	506	2 T14907	trans-cinnamate 4-
12	45	46.9	396	2 T30953	hypothetical prote
13	44	45.8	506	2 F85016	probable RING zinc
14	43	44.8	166	2 T32701	hypothetical prote
15	43	44.8	529	2 T00677	hypothetical prote
16	43	44.8	569	2 T52056	vacuolar protein s
17	43	44.8	569	2 T00445	vacuolar protein-s
18	42	43.8	159	2 C95967	lactoylglytation 1
19	42	43.8	482	2 AE3381	ATP-dependent RNA
20	41	42.7	466	2 S19365	hypothetical prote
21	41	42.7	469	2 C70357	hypothetical prote
22	41	42.7	837	2 B82932	preprotein translo
23	41	42.7	2160	2 T20241	hypothetical prote
24	40	41.7	51	2 G81188	hypothetical prote
25	40	41.7	176	2 G64049	inorganic pyrophos
26	40	41.7	220	2 T08628	hypothetical prote
27	40	41.7	262	2 E75516	conserved hypothet
28	40	41.7	292	1 B65040	YifB protein - Esc
29	40	41.7	292	2 AF0834	conserved hypothet

ALIGNMENTS

RESULT 1

S03849

ribonucleoprotein La - bovine

N;Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B

C;Species: Bos primigenius taurus (cattle)

C;Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C;Accession: S03849

R;Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.

Nucleic Acids Res. 17, 2233-2244, 1989

A;Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequence

A;Reference number: S03848; MUID:89202037; PMID:2468131

A;Accession: S03849

A;Molecule type: mRNA

A;Residues: 1-404 <CHA>

A;Cross-references: EMBL:X13698; NID:g755; PIDN:CAA31986.1; PID:g756

A;Note: part of this sequence was confirmed by protein sequencing

C;Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.

C;Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C;Keywords: blocked amino end; phosphoprotein; RNA binding

F;112-178/Domain: ribonucleoprotein repeat homology <RPM>

F;113-118/Region: RNA-binding RNP2 motif

F;151-158/Region: RNA-binding RNP1 motif

F;228-404/Domain: phosphorylated #status predicted <PHY>

Query Match 90.6%; Score 87; DB 1; Length 404;

Best Local Similarity 88.9%; Pred.No. 6.8e-07;

Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQICQIEYFGDF 18

|||||:|||||

Db 11 AALEAKICHQIEYFGDF 28

RESULT 2

A31888

ribonucleoprotein La - human

N;Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome an

C;Species: Homo sapiens (man)

C;Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C;Accession: A31888; S03848; A22956; A61051; S11013; I55553; I70205; I70206; A31273

R;Chambers, J.C.; Kenan, D.; Martin, B.J.; Keene, J.D.

J. Biol. Chem. 263, 18043-18051, 1988

A;Title: Genomic structure and amino acid sequence domains of the human La autoanti

A;Reference number: A31888; MUID:89053970; PMID:3192525

A;Accession: A31888

A;Molecule type: mRNA

A;Residues: 1-408 <CHA>

A;Cross-references: GB:J04205; NID:g178686; PIDN:AAA51885.1; PID:g178687

R;Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.

Nucleic Acids Res. 17, 2233-2244, 1989

A;Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequence

A:Reference number: S03848; MUID:89202037; PMID:2468131

A:Accession: S03848

A:Molecule type: mRNA

A:Residues: 1-408 <CH2>

A:Cross-references: EMBL:X13697; NID:g36414; PIDN:CAA31985.1; PID:g36415

R:Chambers, J.C.; Keene, J.D.

Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985

A:Title: Isolation and analysis of cDNA clones expressing human lupus La antigen.

A:Reference number: A22956; MUID:85166283; PMID:3856888

A:Accession: A22956

A:Molecule type: mRNA

A:Residues: 45-97, 'LK' <CH3>

A:Cross-references: GB:J04205

A:Note: This sequence has been revised in reference A31888

R:Nyman, U.; Ringertz, N.R.; Pettersson, I.

Immunol. Lett. 22, 65-72, 1989

A:Title: Demonstration of an amino terminal La epitope recognized by human anti-La sera.

A:Reference number: A61051; MUID:89379261; PMID:2476379

A:Accession: A61051

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-19, 'E', 21-47 <NYM>

R:Sturgess, A.D.; Peterson, M.G.; McNeillage, L.J.; Whittingham, S.; Coppel, R.L.

J. Immunol. 140, 3212-3218, 1988

A:Title: Characteristics and epitope mapping of a cloned human autoantigen La.

A:Reference number: S11013; MUID:88199081; PMID:2452201

A:Accession: S11013

A:Molecule type: mRNA

A:Residues: 'E', 55-287, 'V', 289-408 <STUS>

A:Cross-references: EMBL:M20328; NID:g337456; PIDN:AAA36577.1; PID:g337457

R:Kohsaka, H.; Yamamoto, K.; Fujii, H.; Miura, H.; Miyasaka, N.; Nishioka, K.; Miyamoto, J.

Clin. Invest. 85, 1566-1574, 1990

A:Title: Fine epitope mapping the human SS-B/La protein: Identification of a distinct au

A:Reference number: I55553; MUID:90237237; PMID:1692037

A:Accession: I55553

A>Status: translated from GB/EMBL/DBBJ

A:Molecule type: mRNA

A:Residues: 81-107 <RES>

A:Cross-references: GB:M35261; NID:g338491; PIDN:AAA36652.1; PID:g338495

A:Accession: I70205

A>Status: translated from GB/EMBL/DBBJ

A:Molecule type: mRNA

A:Residues: 174-224 <RE3>

A:Cross-references: GB:M35263; NID:g338492; PIDN:AAA36653.1; PID:g338496

A:Accession: I70206

A>Status: translated from GB/EMBL/DBBJ

A:Molecule type: mRNA

A:Residues: 279-342 <RE3>

A:Cross-references: GB:M35262; NID:g338493; PIDN:AAA36654.1; PID:g338497

C:Comment: This protein associates with a variety of small RNA molecules, most of which

ay act as a transcription termination factor.

C:Genetics:

A:Gene: GDB:SSB

A:Cross-references: GDB:125359; OMIM:109090

A:Map position: 2

A:Introns: 22/3; 57/2; 115/3; 151/3; 185/2; 209/2; 223/3; 264/3; 380/2

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: phosphoprotein; RNA binding

F:112-178/Domain: ribonucleoprotein repeat homology <RRM>

F:113-118/Region: RNA-binding RNP2 motif

F:151-158/Region: RNA-binding RNP1 motif

F:228-408/Domain: phosphorylated #status experimental <PHY>

Query Match 90.6%; Score 87; DB 1; Length 408;

Best Local Similarity 88.9%; Pred. No. 6.9e-07;

Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIQIEYFGDF 18

|||||:|||||

Db 11 AALEAKICHQIEYFGDF 28

RESULT 3

JC1494

ribonucleoprotein La - rat

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B

C:Species: Rattus norvegicus (Norway rat)

C:Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: JC1494; S25145

R:Sensel, I.; Trooster, H.; Bartsch, H.; Schwemmlie, M.; Igloi, G.L.; Bachmann, M.

Gene 126, 265-268, 1993

A:Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection

A:Reference number: JC1494; MUID:93246255; PMID:7916708

A:Accession: JC1494

A:Molecule type: mRNA

A:Residues: 1-415 <SEM>

A:Cross-references: GB:X67859; NID:g55778; PIDN:CAA48043.1; PID:g55779

A:Experimental source: liver

C:Comment: This protein associates with a variety of small RNA molecules, most of wh

ay act as a transcription termination factor.

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: phosphoprotein; RNA binding

F:112-178/Domain: ribonucleoprotein repeat homology <RRM>

F:113-118/Region: RNA-binding RNP2 motif

F:151-158/Region: RNA-binding RNP1 motif

F:227-415/Domain: phosphorylated #status predicted <PHY>

Query Match 90.6%; Score 87; DB 1; Length 415;

Best Local Similarity 88.9%; Pred. No. 7e-07;

Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIQIEYFGDF 18

|||||:|||||

Db 11 AALEAKICHQIEYFGDF 28

RESULT 4

S33817

ribonucleoprotein La.B - African clawed frog

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B

C:Species: Xenopus laevis (African clawed frog)

C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: S33817; S28544

R:Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.

J. Mol. Biol. 231, 196-204, 1993

A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression.

A:Reference number: S33817; MUID:93287095; PMID:8510143

A:Accession: S33817

A:Molecule type: mRNA

A:Residues: 1-427 <SCH>

A:Cross-references: EMBL:X68818; NID:g64875; PIDN:CAA48716.1; PID:g64876

C:Comment: This protein associates with a variety of small RNA molecules, most of wh

ay act as a transcription termination factor.

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: phosphoprotein; RNA binding

F:111-177/Domain: ribonucleoprotein repeat homology <RRM>

F:112-117/Region: RNA-binding RNP2 motif

F:150-157/Region: RNA-binding RNP1 motif

F:227-427/Domain: phosphorylated #status predicted <PHY>

Query Match 70.8%; Score 68; DB 1; Length 427;

Best Local Similarity 73.3%; Pred. No. 0.0012;

Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LEAQICQIQIEYFGD 17

||:|||||

Db 12 LDTKICEQIEYFGD 26

RESULT 5

S33818

ribonucleoprotein La.A - African clawed frog

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B

C:Species: Xenopus laevis (African clawed frog)

C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: S33818; S28545

R;Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.

J. Mol. Biol. 231, 196-204, 1993

A;Title: La proteins from *Xenopus laevis*. cDNA cloning and developmental expression.

A;Reference number: S33817; MUID:93287095; PMID:8510143

A;Accession: S33818

A;Molecule type: mRNA

A;Residues: 1-428 <SCH>

A;Cross-references: EMBL:X68817; NID:964873; PID:CAA48715.1; PID:964874

C;Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.

C;Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C;Keywords: phosphoprotein; RNA binding

F:112-178/Domain: ribonucleoprotein repeat homology <RRM>

F:113-118/Region: RNA-binding RNP2 motif

F:151-158/Region: RNA-binding RNP1 motif

F:228-428/Domain: phosphorylated status predicted <PHY>

Query Match 66.7%; Score 64; DB 1; Length 428;

Best Local Similarity 71.4%; Pred. No. 0.0058;

Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 EAQICQIEYYFGD 17

Db 14 DTKICEQIEYYFGD 27

: :||:|||||||

RESULT 6

A53773

La/SS-B homolog D-la - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000

C;Accession: A53773

R;Bal, C.; Li, Z.; Tolias, P.P.

Mol. Cell. Biol. 14, 5123-5129, 1994

A;Title: Developmental characterization of a *Drosophila* RNA-binding protein homologous to

A;Reference number: A53773; MUID:94309632; PMID:8035794

A;Accession: A53773

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-390 <BAT>

A;Cross-references: GB:U07652; NID:9464019; PID:AAA20518.1; PID:9464020

C;Genetics:

A;Gene: FlyBase:La

A;Cross-references: FlyBase:FBgn0011638

C;Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C;Keywords: leucine zipper; RNA binding

Query Match 54.2%; Score 52; DB 2; Length 390;

Best Local Similarity 64.3%; Pred. No. 0.58;

Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 EAQICQIEYYFGD 17

Db 51 ERAIRQVEYYFGD 64

: :||:|||||||

RESULT 7

A53781

ribonucleoprotein La - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 20-Sep-1999

C;Accession: A53781

R;Yoo, C.J.; Wolin, S.L.

Mol. Cell. Biol. 14, 5412-5424, 1994

A;Title: La proteins from *Drosophila melanogaster* and *Saccharomyces cerevisiae*: a yeast

A;Reference number: A53781; MUID:94309661; PMID:8035818

A;Accession: A53781

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-390 <YOO>

A;Cross-references: GB:L32988; NID:9488469; PID:9488470

C;Genetics:

A;Gene: FlyBase:La

A;Cross-references: FlyBase:FBgn0011638

C;Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C;Keywords: RNA binding

Query Match 54.2%; Score 52; DB 2; Length 390;

Best Local Similarity 64.3%; Pred. No. 0.58;

Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 EAQICQIEYYFGD 17

Db 51 ERAIRQVEYYFGD 64

: :||:|||||||

RESULT 8

H89777

capsular polysaccharide synthesis enzyme Cap5p [imported] - *Staphylococcus aureus*

C;Species: *Staphylococcus aureus*

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C;Accession: H89777

R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: H89777

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-391 <KUR>

A;Cross-references: GB:BA000018; PID:gl3700080; PIDN:BA841379.1; GSPDB:GN00149

A;Experimental source: strain N315

C;Genetics:

A;Gene: cap

C;Superfamily: lipopolysaccharide biosynthesis protein bplD

Query Match 52.1%; Score 50; DB 2; Length 391;

Best Local Similarity 72.7%; Pred. No. 1.3;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 QICQIEYYFG 16

Db 366 RICEAIEYYFG 376

:||:|||||

RESULT 9

T38937

rna binding protein - fission yeast (*Schizosaccharomyces pombe*)

C;Species: *Schizosaccharomyces pombe*

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C;Accession: T38937

R;Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, April 1997

A;Reference number: 221818

A;Accession: T38937

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-298 <BAD>

A;Cross-references: EMBL:Z94864; PIDN:CAB08173.1; GSPDB:GN00066; SPDB:SPAC57A10.10

A;Experimental source: strain 972h-; cosmid c57A10

C;Genetics:

A;Gene: SPDB:SPAC57A10.10c

A;Map position: 1

A;Introns: 72/1

Query Match 51.0%; Score 49; DB 2; Length 298;

Best Local Similarity 50.0%; Pred. No. 1.4;

Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 4 EAQICQIEYYFGD 17

Db 64 EAEVLKQVEYFSD 77

:||:|||||

```

Query Match      49.0%; Score 47; DB 2; Length 506;
Best Local Similarity 50.0%; Pred. No. 5.4;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

dy 5 AQICQIQIYYFGDF 18
   :: | |||
db 215 SRLAQSFYHFGDF 228

```

hypothetical protein C14C6.12 - Caenorhabditis elegans.
Species: Caenorhabditis elegans
Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text
Accession: T32701
David, M.; Wohldmann, P.; Bauer, C.; Antoniou, B.
Submitted to the EMBL Data Library, December 1997
Description: The sequence of C. elegans cosmid C14C6.
Reference number: Z21210
Accession: T32701
Status: preliminary; translated from GB/EMBL/DDBJ

Search completed: April 23, 2003, 13:34:33
Job time : 12.1124 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:34:48 ; Search time 20.0225 Seconds
(Without alignments)
185.234 Million cell updates/sec

Title: US-09-836-073-14

Perfect score: 99
Sequence: 1 AALEAKICHQIEYYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 9297

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	33.3	24	Q9BM09	Q9BM09 spongilla 1
2	29	29.3	16	Q9R963	Q9R963 helicobacte
3	28	28.3	11	Q9UC46	Q9UC46 homo sapien
4	28	28.3	24	Q05616	Q05616 staphylococ
5	27.5	27.8	24	Q38270	Q38270 bacterioph
6	27	27.3	18	Q16028	Q16028 homo sapien
7	27	27.3	24	Q28430	Q28430 gorilla gor
8	27	27.3	24	Q28781	Q28781 Pongo pygma
9	27	27.3	25	Q30249	Q30249 haplochromi
10	26	26.3	8	Q9BY5	Q9BY5 homo sapien
11	26	26.3	8	Q9BFA0	Q9BFA0 macaca mula
12	26	26.3	8	Q9BF99	Q9BF99 mylobates c
13	26	26.3	21	Q41496	Q41496 solanum tub
14	26	26.3	24	Q9TNS7	Q9TNS7 homo sapien
15	26	26.3	24	Q9TNS6	Q9TNS6 homo sapien
16	25	25.3	13	Q9ELV4	Q9ELV4 hepatitis b

17	25	25.3	13	12	Q9ELV3	Q9ELV3 hepatitis b
18	25	25.3	13	12	Q9ELV2	Q9ELV2 hepatitis b
19	25	25.3	13	12	Q9ELV1	Q9ELV1 hepatitis b
20	25	25.3	13	12	Q9ELV0	Q9ELV0 hepatitis b
21	25	25.3	13	12	Q9ELU9	Q9ELU9 hepatitis b
22	25	25.3	13	12	Q9ELU8	Q9ELU8 hepatitis b
23	25	25.3	13	12	Q9ELU6	Q9ELU6 hepatitis b
24	25	25.3	13	12	Q9ELU5	Q9ELU5 hepatitis b
25	25	25.3	13	12	Q9ELU4	Q9ELU4 hepatitis b
26	25	25.3	13	12	Q9ELU3	Q9ELU3 hepatitis b
27	25	25.3	13	12	Q9ELU2	Q9ELU2 hepatitis b
28	25	25.3	13	12	Q9ELU1	Q9ELU1 hepatitis b
29	25	25.3	13	12	Q9ELU0	Q9ELU0 hepatitis b
30	25	25.3	13	12	Q9ELT9	Q9ELT9 hepatitis b
31	25	25.3	13	12	Q9ELT8	Q9ELT8 hepatitis b
32	25	25.3	25	8	Q9TGB8	Q9TGB8 alnus crisp
33	25	25.3	25	8	Q9TGB7	Q9TGB7 alnus glut
34	25	25.3	25	8	Q9TGB6	Q9TGB6 alnus marit
35	25	25.3	25	8	Q9TGB5	Q9TGB5 betula alle
36	25	25.3	25	8	Q9TGB4	Q9TGB4 betula glan
37	25	25.3	25	8	Q9TGB3	Q9TGB3 betula papy
38	25	25.3	25	8	Q9TGB2	Q9TGB2 betula verr
39	25	25.3	25	8	Q9TGB1	Q9TGB1 betula pube
40	25	25.3	25	8	Q9TGB0	Q9TGB0 corylus ave
41	25	25.3	25	8	Q9TGA9	Q9TGA9 corylus col
42	25	25.3	25	8	Q9TGA8	Q9TGA8 corylus cor
43	25	25.3	25	8	Q9TGA7	Q9TGA7 ostrya virg
44	25	25.3	25	8	Q9TGA6	Q9TGA6 quercus rub
45	25	25.3	25	8	Q9TGB3	Q9TGB3 carpinus ca

ALIGNMENTS

RESULT 1

Q9BM09 PRELIMINARY; PRT; 24 AA.
AC Q9BM09; 17, Created
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Gypsy-like reverse transcriptase (Fragment).
OS Spongilla lacustris (Freshwater sponge).
OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;
OC Haplosclerida; Spongillidae; Spongillia.
OX NCBI_TaxID=6055;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSOM-GRT-G7 RETROTRANSPOSOM;
RX MEDLINE=20570504; PubMed=11121049;
RA Arkhipova I., Meselson M.;
RT "Transposable elements in sexual and ancient asexual taxa.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
DR EMBL; AY013997; AAG59969.1; -;
KW RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 24
SQ SEQUENCE 24 AA; 2754 MW; B588512EDE7A251A CRC64;

Query Match 33.3%; Score 33; DB 5; Length 24;
Best Local Similarity 44.4%; Pred. No. 1.4e-02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIEYYF 15

DB 15 VCHGLEFVF 23

RESULT 2

Q9R963 PRELIMINARY; PRT; 16 AA.
ID Q9R963
AC Q9R963;
DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE FpC (Fragment).
 GN FpC.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=210;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=F31;
 RC MEDLINE=98453456; PubMed=9780260;
 RX Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,
 RA "Full-length sequence analysis of the vacA gene from cytotoxic and
 RT noncytotoxic Helicobacter pylori";
 RT J. Infect. Dis. 178:1391-1398 (1998).
 DR EMBL: AF049623; AAD04263.1; -;
 FT NON_TER 1
 FT NON_TER 16
 SQ SEQUENCE 16 AA; 1938 MW; 3C33D403AEE61428 CRC64;
 Query Match 29.3%; Score 29; DB 2; Length 16;
 Best Local Similarity 55.6%; Pred. No. 4.4e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 10 QIEYFGDF 18
 I:||||:
 Db 7 QVEYAFNF 15
 RESULT 3
 Q9UC46 PRELIMINARY; PRT; 11 AA.
 AC Q9UC46;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE Neutrophil inhibitor peptide, NIP-POLYMORPHONUCLEAR neutrophil
 DE inhibitor peptide.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=96326114; PubMed=8703476;
 RA Cooper J.A.Jr., Culbreth R.R.;
 RT "Characterization of a neutrophil inhibitor peptide harvested from
 RT human bronchial lavage: homology to Influenza A nucleoprotein.";
 RL Am. J. Respir. Cell Mol. Biol. 15:207-215 (1996)
 SQ SEQUENCE 11 AA; 1262 MW; 951A1C3279C9DB45 CRC64;
 Query Match 28.3%; Score 28; DB 4; Length 11;
 Best Local Similarity 80.0%; Pred. No. 4.5e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 13 YIFGD 17
 I:||||:
 Db 5 YIFGD 9
 RESULT 4
 Q05616 PRELIMINARY; PRT; 24 AA.
 AC Q05616;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE 3-dehydroquinase synthase (EC 4.6.1.3) (Fragment).
 GN AROB.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8325-4;
 RX MEDLINE=93381456; PubMed=8371108;
 RA O'Connell C.M., Pattee P., Foster T.J.;
 RT "Sequence and mapping of the araA gene of Staphylococcus aureus 8325-
 RT 4.";
 RL J. Gen. Microbiol. 139:1449-1460 (1993).
 CC -!- CATALYTIC ACTIVITY: 7-PHOSPHO-3-DEOXY-ARABINO-HEPTULOSONATE - 3-
 CC DEHYDROQUINATE + ORTHOPHOSPHATE.
 CC -!- PATHWAY: SECOND STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE
 CC AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
 DR EMBL: L05004; AAA71896.1; -;
 DR Aromatic amino acid biosynthesis; Lyase.
 KW NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 24 AA; 2857 MW; 1C4C8D20F7A4856E CRC64;
 Query Match 28.3%; Score 28; DB 2; Length 24;
 Best Local Similarity 50.0%; Pred. No. 9.8e+02;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 8 CHOIEYF 15
 I:|:|:
 Db 16 CEQLKTYF 23
 RESULT 5
 Q38270 PRELIMINARY; PRT; 24 AA.
 AC Q38270;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created).
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Bacteriophage lambda mutant truncated regulatory protein (PyrI) gene
 DE from E.coli, 3' end (Fragment).
 OS Bacteriophage lambda.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
 OC Lambda-like viruses.
 OX NCBI_TaxID=10710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86143826; PubMed=3912513;
 RA Cunin R., Jacobs A., Charlier D., Crabeel M., Herve G., Glansdorff N.,
 RA Pierard A.;
 RT "Structure-function relationship in allosteric aspartate
 RT carbamoyltransferase from Escherichia coli: I. Primary structure of a
 RT pyrI gene encoding a modified regulatory subunit.";
 RL J. Mol. Biol. 186:707-713 (1985).
 DR EMBL: M28579; AAA3252.1; -;
 DR HSP; P00478; BATC.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 24 AA; 2905 MW; FD9349DF6F8159D0 CRC64;
 Query Match 27.8%; Score 27.5; DB 9; Length 24;
 Best Local Similarity 46.2%; Pred. No. 1.2e+03;
 Matches 6; Conservative 3; Mismatches 3; Indels 1; Gaps 1;
 QY 2 ALEAKICHOIEY 14
 I:|:|:|:
 Db 8 ALKCYCEK-EFY 19
 RESULT 6
 Q16028 PRELIMINARY; PRT; 18 AA.
 ID Q16028
 AC Q16028;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
 DE OCRL-1 protein (Fragment).


```

GN OCRL-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93279398; PubMed=8504307;
RA Leahy A.M., Charnas L.R., Nussbaum R.L.;
RT "Nonsense mutations in the OCRL-1 gene in patients with the
RT oculocephrorenal syndrome of Lowe.";
RL Hum. Mol. Genet. 2:461-463(1993).
DR EMBL; S62084; AAD13933.1; -
FT NON_TER 1
SQ SEQUENCE 18 AA; 2055 MW; 61356C825F7AF33D CRC64;

Query Match 27.3%; Score 27; DB 4; Length 18;
Best Local Similarity 36.4%; Pred. No. 1.4e+03;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 AALEAKICHQI 11
Db 1 SAYDPRICRQL 11

RESULT 7
Q28430 PRELIMINARY; PRT; 24 AA.
AC Q28430;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Relaxin (Fragment).
GN RLX.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RG YK;
RA EVANS B.B.A., Fu P., Tregear G.G.W.;
RT "Characterisation of primate relaxin genes.";
RL J. Mol. Endocrinol. 0:0-0(1993).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
DR EMBL; 227228; CA81742.1; -
DR HSSP; P04090; 6RLX.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PROSITE; PS00262; INSULIN; 1.
FT NON_TER 1
SQ SEQUENCE 24 AA; 2673 MW; 5615D4B92C1DFD24 CRC64;

Query Match 27.3%; Score 27; DB 6; Length 24;
Best Local Similarity 55.6%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AALEAKICHQI 11
Db 1 SAYDPRICRQL 11

RESULT 8
Q28781 PRELIMINARY; PRT; 24 AA.
AC Q28781;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Relaxin (Fragment).
GN RLX.
OS Pongo pygmaeus (Orangutan).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RG YK 287;
RA EVANS B.B.A., Fu P., Tregear G.G.W.;
RT "Characterisation of primate relaxin genes.";
RL J. Mol. Endocrinol. 0:0-0(1993).
CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL; 227229; CA81743.1; -
DR HSSP; P04090; 6RLX.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PROSITE; PS00262; INSULIN; 1.
FT NON_TER 1
SQ SEQUENCE 24 AA; 2673 MW; 5615D4B92C1DFD24 CRC64;

Query Match 27.3%; Score 27; DB 6; Length 24;
Best Local Similarity 55.6%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AALEAKICHQI 9
Db 4 SALANKCCH 12

RESULT 9
Q90249 PRELIMINARY; PRT; 25 AA.
AC Q90249;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Gonadotropin-releasing hormone type II receptor (Fragment).
OS Haplochromis burtoni (Burton's mouthbrooder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidel;
OC Cichlidae; Astatotilapia.
OX NCBI_TaxID=8153;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21216506; PubMed=11316736;
RA Robison R.R., White R.B., Illing N., Troskie B.E., Morley M.,
RA Millar R.P., Fernald R.D.;
RT "Gonadotropin-releasing hormone receptor in the teleost Haplochromis
RT burtoni: structure, location, and function.";
RL Endocrinology 142:1737-1743(2001).
DR EMBL; AF356598; AAK52751.1; -
KW Receptor.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 2692 MW; 3677F81D5337E8C4 CRC64;

Query Match 27.3%; Score 27; DB 13; Length 25;
Best Local Similarity 36.4%; Pred. No. 1.5e+03;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LEAKICHQIEY 13
Db 4 LEGKVSHSLTH 14

RESULT 10
Q9BYI5 PRELIMINARY; PRT; 8 AA.
AC Q9BYI5;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CAMP responsive element modulator (Fragment).

```

```

GN CREM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011664; AAG47575.1; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 1006 MW; DF02C331EEAB572A CRC64;

Query Match 26.3%; Score 26; DB 4; Length 8;
Best Local Similarity 60.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHOIE 12
DB 4 CHKVE 8

RESULT 11
Q9BFA0
ID Q9BFA0 PRELIMINARY; PRT; 8 AA.
AC Q9BFA0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CAMP responsive element moderator (Fragment).
GN CREM.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011661; AAG47572.1; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 1006 MW; DF02C331EEAB572A CRC64;

Query Match 26.3%; Score 26; DB 6; Length 8;
Best Local Similarity 60.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHOIE 12
DB 4 CHKVE 8

RESULT 12
Q9BF99
ID Q9BF99 PRELIMINARY; PRT; 8 AA.
AC Q9BF99;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CAMP responsive element moderator (Fragment).
GN CREM.
OS Hylobates concolor (crested gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=29089;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011662; AAG47573.1; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 1006 MW; DF02C331EEAB572A CRC64;

Query Match 26.3%; Score 26; DB 6; Length 8;
Best Local Similarity 60.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHOIE 12
DB 4 CHKVE 8

RESULT 13
Q41496
ID Q41496 PRELIMINARY; PRT; 21 AA.
AC Q41496;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Truncated proteinase inhibitor I.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. RUSSET BURBANK;
RA Lee J.S., Yang Y.S.;
RT "Nucleotide Sequence of a Truncated Proteinase Inhibitor I Gene of
RT Potato.";
RL Singmul Hakhoe Chi 33:303-307(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. RUSSET BURBANK;
RA Lee J.S.;
RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z12819; CAA78280.1; -.
SQ SEQUENCE 21 AA; 2608 MW; 21B4922FFB25750F CRC64;

Query Match 26.3%; Score 26; DB 10; Length 21;
Best Local Similarity 33.3%; Pred. No. 1.9e+03;
Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 LEAKICHQIEYY 14
DB 1 MESKFAHIIIVFF 12

RESULT 14
Q9TNS7
ID Q9TNS7 PRELIMINARY; PRT; 24 AA.
AC Q9TNS7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 14, Last annotation update)
DE Human leucocyte antigen beta chain DR molecule HLA-DRB1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93107319; PubMed=1469092;
RA Weyand C.M., Hicok K.C., Hunder G.G., Goronzy J.J.;
RT "The HLA-DRB1 locus as a genetic component in giant cell arteritis.

```

RT Mapping of a disease-linked sequence motif to the antigen binding site

RT of the HLA-DR molecule."

RL J. Clin. Invest. 90:2355-2361(1992).

KW MHC.

SQ SEQUENCE 24 AA; 3042 MW; 46FF753670C7A760 CRC64;

Query Match 26.3%; Score 26; DB 7; Length 24;

Best Local Similarity 20.0%; Pred. No. 2.1e+03;

Matches 2; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIEYF 15

:: |::

Db 3 QVKHEFDYF 12

RESULT 15

Q9TNS6

ID Q9TNS6 PRELIMINARY; PRT; 24 AA.

AC Q9TNS6;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE Human leucocyte antigen beta chain DR molecule HLA-DRB1 (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93107319; PubMed=1469092;

RA Weyand C.M., Hicok K.C., Hunder G.G., Goronzy J.J.;

RT "The HLA-DRB1 locus as a genetic component in giant cell arteritis.

RT Mapping of a disease-linked sequence motif to the antigen binding site

RT of the HLA-DR molecule."

RL J. Clin. Invest. 90:2355-2361(1992).

KW MHC.

SQ SEQUENCE 24 AA; 3112 MW; 5C4F753667F7A760 CRC64;

Query Match 26.3%; Score 26; DB 7; Length 24;

Best Local Similarity 20.0%; Pred. No. 2.1e+03;

Matches 2; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIEYF 15

:: |::

Db 3 QVKHEFDYF 12

Search completed: April 23, 2003, 13:47:15

Job time : 20.0225 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003. Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:33:13 ; Search time 4.65169 Seconds
(without alignments)
160.495 Million cell updates/sec

Title: US-09-836-073-14

Perfect score: 99
Sequence: 1 AALEAKICHOIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1520

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	27	27.3	19	1 HBB2_UROHA	P18992 uromastix h
2	24	24.2	11	1 CXLL1_CONMR	P58807 conus marmo
3	24	24.2	12	1 TIN4_HOPTI	P82652 holoibatrac
4	24	24.2	13	1 CXL4_CONMR	P58810 conus marmo
5	24	24.2	16	1 MLB_SQUAC	P01207 squalus sca
6	24	24.2	18	1 MLB_SCYCA	P01206 scyllorhinu
7	24	24.2	25	1 CXOB_CONMA	P03485 conus magus
8	22	22.2	21	1 BTX_ATRBI	P80163 attractaspis
9	22	22.2	21	1 SRTD_ATREN	P13211 attractaspis
10	22	22.2	23	1 XYC1_ACIGB	P46365 acinetobact
11	22	22.2	25	1 ANDT_ANDAU	P56684 androctonus
12	21	21.2	17	1 TPIS_PINPS	P81666 pinus pinas
13	21	21.2	23	1 NUO5_SOLTU	P80262 solanum tub
14	21	21.2	25	1 SMBP_RAT	P80968 rattus norv
15	20	20.2	10	1 TKNB_RANRI	P29135 rana ridibu
16	20	20.2	14	1 CXAL_CONCN	P56973 conus conso
17	20	20.2	15	1 PC20_BRANA	P81096 brassica na
18	20	20.2	16	1 CXA2_CONMA	P56636 conus magus
19	20	20.2	18	1 OBP_LYMDI	P34173 lymanthria d
20	20	20.2	18	1 SFAH_HELAN	P81098 helianthus
21	20	20.2	20	1 COG4_CHIOP	P34156 chionoecete
22	20	20.2	20	1 CS21_STRTR	P81621 streptococc
23	20	20.2	20	1 FIBB_FELCA	P14469 felis silve
24	20	20.2	20	1 TL18_SPIOL	P84536 spinacia ol
25	20	20.2	21	1 NDK_CANAL	Q9ur66 candida alb
26	20	20.2	22	1 LPL_TRIWA	P24335 trimeresuru
27	20	20.2	22	1 LP2_TRIWA	P58930 trimeresuru
28	20	20.2	24	1 POQA_ACICA	P27532 acinetobact
29	20	20.2	24	1 POQA_PSEFL	P55171 pseudomonas
30	19.5	19.7	20	1 COXN_THUOB	P80980 thunnus obo
31	19.5	19.7	23	1 UDP_LACCA	P19662 lactobacill
32	19	19.2	8	1 AL17_CARMA	P81820 carcinus ma
33	19	19.2	9	1 AL11_CARMA	P81814 carcinus ma

34 19 19.2 10 1 GON2_CHEPR P80678 chelyosoma
35 19 19.2 10 1 TRNB_ONCMY P88500 oncorhynchu
36 19 19.2 11 1 TIN4_HOPTI P82654 holoibatrac
37 19 19.2 12 1 NO40_SESRO O24369 sesbania ro
38 19 19.2 12 1 TIN3_HOPTI P82653 holoibatrac
39 19 19.2 14 1 ADF_TENMO P82965 tenebrio mo
40 19 19.2 14 1 MAST_VESLE P01514 vespula lew
41 19 19.2 17 1 B29K_PORGI P81784 porphyronon
42 19 19.2 19 1 FIBB_VULVU P14482 vulpes vulp
43 19 19.2 21 1 PEDB_HYDAT P80577 hydra atten
44 19 19.2 21 1 TRNC_CARAU P25421 carassius a
45 19 19.2 25 1 PRLA_ACHLY P27459 achromobact

ALIGNMENTS

RESULT 1
HBB2_UROHA
ID HBB2_UROHA STANDARD: PRT; 19 AA.
AC P18992;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hemoglobin beta-2 chain (Fragment).
OS Uromastix hardwickii (Indian spiny-tailed lizard).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Uromastycinae;
OC Uromastix.
OX NCBI_TaxID=40250;
RN [1]
RP SEQUENCE.
RX MEDLINE=84029159; PubMed=6628672;
RA Naqvi S., Zaidi Z.H., von Bahr-Lindstroem H., Carlquist M.,
RA Joernvall H.;
RT "Characterization of hemoglobin from the lizard Uromastix
hardwickii.";
RL FEBS Lett. 162:290-295(1983).
CC -!- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CC -!- VARIOUS PERIPHERAL TISSUES.
CC -!- SUBUNIT: HETEROETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -!- TISSUE SPECIFICITY: RED BLOOD CELLS.
DR PIR: A05305; A05305.
DR InterPro: IPR000971; Globin.
DR PROSITE: PS01033; GLOBIN; PARTIAL.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT NON_TER 1
FT NON_TER 19
SQ SEQUENCE 19 AA; 1914 MW; C40AD8EA30019057 CRC64;

Query Match 27.3%; Score 27; DB 1; Length 19;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0.

QY 14 YFGDF 18

Db 1 FFGDF 5

RESULT 2

CXLL1_CONMR STANDARD: PRT; 11 AA.
ID CXLL1_CONMR
AC P58807;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda-conotoxin CMrVIA.
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]

```

RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE-Venom;
RX MEDLINE-20564325; PubMed-10988292;
RA Balaji R.A., Ontake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus";
RL J. Biol. Chem. 275:39516-39522(2000).
CC -|- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -|- MASS SPECTROMETRY: MW-1237.93; MW ERR-0.21; METHOD-Electrospray.
CC -|- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KW Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 2 11
FT DISULFID 3 8
FT MOD_RES 10 10 HYDROXYLATION.
FT MOD_RES 11 AA; 1226 MW; 277AAC60B7232B58 CRC64;
SQ SEQUENCE 11 AA; 1226 MW; 277AAC60B7232B58 CRC64;

Query Match 24.2%; Score 24; DB 1; Length 11;
Best Local Similarity 75.0%; Pred. NO. 4.3e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 KICH 9
DB 6 KLCH 9

RESULT 3
TIN2_HOPTI
ID TIN2_HOPTI STANDARD; PRT; 12 AA.
AC P82652;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tigerin-2
OS Hoplobatrachus tigerinus (Indian bull frog) (Rana tigerina).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae;
OC Hoplobatrachus.
OX NCBI_TaxID=103373;
[1]
RN SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND DISULFIDE BONDS.
RP TISSUE-Skin;
RX PubMed-11031261;
RA Purna Sai K., Jagannadham M.V., Vairamani M., Raju N.P.,
RA Devi A.S., Nagaraj R., Sitaram N.;
RT "Tigerinins: novel antimicrobial peptides from the Indian frog Rana
RT tigerina.";
RL J. Biol. Chem. 276:2701-2707(2001).
CC -|- FUNCTION: ANTIBACTERIAL ACTIVITY AGAINST B.SUBTILIS, E.COLI,
CC S.AUREUS, M.LUTEUS, P.PUTIDA AND S.CEREVISIAE.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: SKIN.
CC -|- MASS SPECTROMETRY: MW-1368; METHOD-MALDI.
KW Amphibian skin; Antibiotic; Amidation.
FT DISULFID 3 11
FT MOD_RES 12 12 AMIDATION.
FT MOD_RES 12 12
SQ SEQUENCE 12 AA; 1369 MW; C2627D8EFD37605D CRC64;

Query Match 24.2%; Score 24; DB 1; Length 12;
Best Local Similarity 50.0%; Pred. NO. 4.7e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALEAKICH 9
DB 5 AIPLPICH 12

RESULT 4
CX14_CONMR
ID CX14_CONMR STANDARD; PRT; 13 AA.

```

```

AC P58910;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda/chi-conotoxin MrIB (Chi-MrIB).
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
[1]
RN SEQUENCE, SYNTHESIS, AND STRUCTURE BY NMR.
RP TISSUE-Venom;
RX MEDLINE-21419681; PubMed-11528421;
RA Sharpe I.A., Gehrmann J., Loughnan M.L., Thomas L., Adams D.A.,
RA Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.,
RA Lewis R.J.;
RT "Two new classes of conopeptides inhibit the alpha1-adrenoceptor and
RT noradrenaline transporter.";
RL Nat. Neurosci. 4:902-907(2001).
CC -|- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -|- PTM: Exists in two forms, due to cis-trans isomerization at His-
CC 11-Hyp-12.
CC -|- MASS SPECTROMETRY: MW-1393.52; METHOD-Electrospray.
CC -|- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
DR PDB; 1IEO; 03-APR-02.
KW Neurotoxin; Toxin; Hydroxylation; 3D-structure.
FT DISULFID 4 13
FT DISULFID 5 10
FT MOD_RES 12 12 HYDROXYLATION.
FT MOD_RES 13 AA; 1382 MW; 277AAC376EAD2B58 CRC64;
SQ SEQUENCE 13 AA; 1382 MW; 277AAC376EAD2B58 CRC64;

Query Match 24.2%; Score 24; DB 1; Length 13;
Best Local Similarity 75.0%; Pred. NO. 5.1e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 KICH 9
DB 8 KLCH 11

RESULT 5
MLB_SQUAC
ID MLB_SQUAC STANDARD; PRT; 16 AA.
AC P01207;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanotropin beta.
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalia; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
[1]
RN SEQUENCE.
RP MEDLINE-75127390; PubMed-4375978;
RX Bennett H.P.J., Lowry P.J., McMartin C., Scott A.P.;
RA "Structural studies of alpha-melanocyte-stimulating hormone and a
RT novel beta-melanocyte-stimulating hormone from the neurointermediate
RT lobe of the pituitary of the dogfish Squalus acanthias.";
RL Biochem. J. 141:439-444(1974).
CC -|- SIMILARITY: BELONGS TO THE POMC FAMILY.
DR PIR; A01471; MTDFBS.
KW Hormone.
SQ SEQUENCE 16 AA; 1939 MW; 993AF43C8A6A0A0E CRC64;

Query Match 24.2%; Score 24; DB 1; Length 16;
Best Local Similarity 57.1%; Pred. NO. 6.2e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 12 EYVFGDF 18
DB 11

```

```

Db      4  DYKFGHF 10

RESULT 6
MLB_SCYCA
ID  MLB_SCYCA  STANDARD;  PRT;  18 AA.
AC  P01206;
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Melanotropin beta (Beta-MSH).
OS  Scyllorhinus canicula (Spotted dogfish). (Spotted catshark).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC  Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC  Scyllorhinidae; Scyllorhinus.
OX  NCBI_TaxID=7830;
RN  [1]
RP  SEQUENCE.
RX  MEDLINE=75113445; PubMed=4452470;
RA  Love R.M., Pickering B.T.;
RT  "A beta-MSH in the pituitary gland of the spotted dogfish
RT  (Scyllorhinus canicula): isolation and structure.";
RL  Gen. Comp. Endocrinol. 24:398-404(1974).
CC  -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
DR  PIR; A01470; MTFBFC.
KW  Hormone.
SQ  SEQUENCE 18 AA; 2195 MW; 353C93A30C314DF4 CRC64;

Query Match 24.2%; Score 24; DB 1; Length 18;
Best Local Similarity 44.4%; Pred. No. 7e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 QIEYFGDF 18
    : : : :
Db 2 ZIBYKMGHF 10

RESULT 7
CXOB_CONMA
ID  CXOB_CONMA  STANDARD;  PRT;  25 AA.
AC  P05485;
DT  01-NOV-1988 (Rel. 09, Created)
DT  01-NOV-1988 (Rel. 09, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Omega-conotoxin MVIIB (SNK-159).
OS  Conus magus (Magus cone).
OC  Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC  Neogastropoda; Conoidea; Conidae; Conus.
OX  NCBI_TaxID=6492;
RN  [1]
RP  SEQUENCE.
RX  MEDLINE=87299637; PubMed=2441741;
RA  Olivera B.M., Cruz L.J., de Santos V., Lecheminant G.W., Griffin D.,
RA  Zeikus R.D., McIntosh J.M., Galyean R., Varga J., Gray W.R.,
RA  Rivier J.E.;
RT  "Neuronal calcium channel antagonists. Discrimination between calcium
RT  channel subtypes using omega-conotoxin from Conus magus venom.";
RL  Biochemistry 26:2086-2090(1987).
CC  -1- FUNCTION: Omega-conotoxins act at presynaptic membranes, they bind
CC  and block voltage-sensitive calcium channels (VSCC).
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC  -1- SIMILARITY: BELONGS TO THE O-SUPERFAMILY OF CONOTOXINS. OMEGA-TYPE
CC  FAMILY.
DR  PIR; B34115; B34115.
DR  PIR; JH0701; JH0701.
DR  HSP; P05484; IMVI.
KW  Presynaptic neurotoxin; Toxin; Calcium channel inhibitor;
AMidation.
FT  DISULFID 1 16
FT  DISULFID 8 20
FT  DISULFID 15 25
FT  MOD_RES 25 25
AMIDATION.

SQ  SEQUENCE 25 AA; 2626 MW; E4B9CE5EFAA3734D CRC64;

Query Match 24.2%; Score 24; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 9.6e+02;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 CHOIEY 13
    : : :
Db 8 CHRTSY 13

RESULT 8
BTX_ATRBI
ID  BTX_ATRBI  STANDARD;  PRT;  21 AA.
AC  P80163;
DT  01-APR-1993 (Rel. 25, Created)
DT  01-APR-1993 (Rel. 25, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Bibtrotoxin (BTX).
OS  Atractaspis bibroni (Southern bibron's) (Stiletto snake).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC  Atractaspididae; Atractaspis.
OX  NCBI_TaxID=8601;
RN  [1]
RP  SEQUENCE.
RX  TISSUE=Venom;
RX  MEDLINE=93106214; PubMed=8416802;
RA  Becker A., Dowdle E.B., Hechler U., Kausser K., Donner P.,
RA  Schleuning W.-D.;
RT  "Bibtrotoxin, a novel member of the endothelin/sarafotoxin peptide
RT  family, from the venom of the burrowing asp Atractaspis bibroni.";
RL  FEBS Lett. 315:100-103(1993).
CC  -1- FUNCTION: VASOCONSTRICTOR ACTIVITY. THESE TOXINS CAUSE CARDIAC
CC  ARREST PROBABLY AS A RESULT OF CORONARY VASOSPASM.
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- SIMILARITY: BELONGS TO THE ENDOTHELIN/SARAFOTOXIN FAMILY.
DR  PIR; S27039; S27039.
DR  HSP; P13208; LSRE.
DR  InterPro; IPR001928; Endothlin_tox.
DR  InterPro; IPR003642; Sara/bib_toxin.
DR  Pfam; PF00322; endothelin_1.
DR  PRINTS; PR00365; ENDOTHELIN.
DR  ProDom; PD004740; Sara/bib_toxin; 1.
DR  SMART; SM00272; END; 1.
DR  PROSITE; PS00270; ENDOTHELIN; 1.
KW  Vasoconstrictor; Toxin.
FT  DISULFID 1 15
FT  DISULFID 3 11
FT  DISULFID 11 15
BY SIMILARITY.
BY SIMILARITY.
SQ  SEQUENCE 21 AA; 2511 MW; 83A5DFB81D036AE2 CRC64;

Query Match 22.2%; Score 22; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQ 10
    : : :
Db 15 CHQ 17

RESULT 9
SRTD_ATREN
ID  SRTD_ATREN  STANDARD;  PRT;  21 AA.
AC  P13211;
DT  01-JAN-1990 (Rel. 13, Created)
DT  01-JAN-1990 (Rel. 13, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Sarafotoxin-D (S6D) (SRTX-D).
OS  Atractaspis engaddensis (Israeli burrowing asp).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC  Atractaspididae; Atractaspis.
OX  NCBI_TaxID=8600;

```

```

RN [1]
RP SEQUENCE.
RC TISSUE-Venom:
RX MEDLINE-90033283; PubMed-2509240;
RA Boelich A., Wollberg Z., Fleming G.;
RT "SRTX-d, a new native peptide of the endothelin/sarafotoxin family.";
RL FEBS Lett. 256:1-3(1989).
CC -1- FUNCTION: VASOCONSTRICTOR ACTIVITY. THESE TOXINS CAUSE CARDIAC
CC -1- ARREST PROBABLY AS A RESULT OF CORONARY VASOSPASM.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: LD50 IS 0.35 MG/KG BY INTRAVENOUS INJECTION.
CC -1- SIMILARITY: BELONGS TO THE ENDOTHELIN/SARAFOTOXIN FAMILY.
DR PIR; S06145; S06145.
DR HSSP; P13208; 1SRB.
DR InterPro; IPR001928; Endothln_tox.
DR InterPro; IPR003642; Sara/bib_toxin.
DR Pfam; PF00322; endothelin; 1.
DR PRINTS; P00365; ENDOTHELIN.
DR PRODOM; PD004740; Sara/bib_toxin; 1.
DR SMART; SM00272; END; 1.
DR PROSITE; PS00270; ENDOTHELIN; 1.
KW Vasoconstrictor; Toxin.
FT DISULFID 1 15 BY SIMILARITY.
FT DISULFID 3 11 BY SIMILARITY.
SQ SEQUENCE 21 AA; 2596 MW; 83A8A04D1D536AE2 CRC64;

Query Match 22.2%; Score 22; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQ 10
DB 15 CHQ 17

RESULT 10
XYCL ACIGB STANDARD; PRT; 23 AA.
AC P46365;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Benzaldehyde dehydrogenase [NAD+] I (EC 1.2.1.28) (Fragment).
OS Acinetobacter genomosp. 11.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=106649;

SEQUENCE.
RP STRAIN-ATCC 11171 / NCIB 8250 / CIP 63.46 / B94;
RX MEDLINE-91113163; PubMed-1989592;
RA Chalmers R.M., Keen J.N., Fawson C.A.;
RT "Comparison of benzyl alcohol dehydrogenases and benzaldehyde
RT dehydrogenases from the benzyl alcohol and mandelate pathways in
RT Acinetobacter calcoaceticus and from the TOL-plasmid-encoded toluene
RT pathway in Pseudomonas putida. N-terminal amino acid sequences, amino
RT acid compositions and immunological cross-reactions.";
RL Biochem J 273:99-107(1991).
CC -1- CATALYTIC ACTIVITY: Benzaldehyde + NAD(+) + H(2)O = benzoate +
CC NADH.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; PARTIAL.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; PARTIAL.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2750 MW; 50496D4930DF9E40 CRC64;

Query Match 22.2%; Score 22; DB 1; Length 23;
Best Local Similarity 40.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

```

```

QY 3 LEAKICHOI--EYFF 15
DB 3 IQTKIIEQIWKHF 17

RESULT 11
ANDT ANDAU STANDARD; PRT; 25 AA.
ID ANDT ANDAU STANDARD; PRT; 25 AA.
AC P56684; P81616;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Androctonin.
OS Androctonus australis hector (Sahara scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Butchoidea; Butthidae; Androctonus.
OX NCBI_TaxID=70175;
RN [1]
RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
RC TISSUE-Hemolymph;
RX MEDLINE-97094646; PubMed-8939880;
RA Ehret-Sabatier L., Loew D., Goyffon M., Fehlbauer P., Hoffmann J.A.,
RA van Dorsselaer A., Bulet P.;
RT "Characterization of novel cysteine-rich antimicrobial peptides from
RT scorpion blood.";
RL J. Biol. Chem. 271:29537-29544(1996).
RN [2]
RP SYNTHESIS OF D-AMINO ACID ENANTIOMER, AND CHARACTERIZATION.
RX MEDLINE-20115101; PubMed-10642525;
RA Hetru C., Letellier L., Oren Z., Hoffmann J.A., Shai Y.;
RT "Androctonin, a hydrophilic disulphide-bridged non-haemolytic
RT anti-microbial peptide: a plausible mode of action.";
RL Biochem. J. 345:653-664(2000).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE-20025109; PubMed-10563585;
RA Mandard N., Sy D., Maufrais C., Bonmatin J.M., Bulet P., Hetru C.,
RA Vovelle F.;
RT "Androctonin, a novel antimicrobial peptide from scorpion Androctonus
RT australis: solution structure and molecular dynamics simulations in
RT the presence of a lipid monolayer.";
RL J. Biomol. Struct. Dyn. 17:367-380(1999).
CC -1- FUNCTION: ACTIVE AGAINST BOTH BACTERIA (GRAM-POSITIVE AND GRAM-
CC NEGATIVE) AND FILAMENTOUS FUNGI. ACTS ON THE MEMBRANE OF THE
CC BACTERIAL CELLS. IT DESTABILIZE A MEMBRANE BY MODIFYING ITS
CC PROPERTIES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MASS SPECTROMETRY: MW=3076.7; METHOD=Electrospray.
DR PDB; 1C26; 12-JAN-00.
KW Antibiotic; Fungicide; 3D-structure.
FT DISULFID 4 20
FT DISULFID 10 16
SQ SEQUENCE 25 AA; 3081 MW; D4183D6EA5AA50AA CRC64;

Query Match 22.2%; Score 22; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIE 12
DB 3 VCRQIK 8

RESULT 12
TPIS PINPS STANDARD; PRT; 17 AA.
ID TPIS PINPS STANDARD; PRT; 17 AA.
AC P81666;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Triosephosphate isomerase, cytosolic (EC 5.3.1.1) (TIM) (Fragments).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

OC Spermatoxipha; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RP SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
proteins";
RL Electrophoresis 20:1098-1108(1999).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glycerone
phosphate.
CC -1- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- INDUCTION: BY WATER STRESS.
CC -1- MISCELLANEOUS: IN PLANTS, THERE ARE TWO TYPES OF TPIS, CYTOSOLIC
AND PLASTID.
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN
(SPOT N139) IS: 5.9, ITS MW IS: 24 kDa.
CC -1- SIMILARITY: BELONGS TO THE TRIOSPHOSPHATE ISOMERASE FAMILY.
DR InterPro: IPR000652; Triophos.ismrse.
DR PROSITE: PS00171; TIM; PARTIAL.
KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
KW Pentose shunt.
FT NON_TER 1 1
FT NON_CONS 9 10
FT NON_TER 17 17
SQ SEQUENCE 17 AA; 2053 MW; 183DB41757AF13CB CRC64;

Query Match 21.2%; Score 21; DB 1; Length 17;
Best Local Similarity 20.0%; Pred. No. 2e+03;
Matches 2; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 7 ICHOIEYFVG 16
Db 4 VCIEQLFFVG 13

RESULT 13
NUO5_SOLTU STANDARD; PRT; 23 AA.
AC P80262;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase 29 kDa subunit (EC 1.6.5.3)
DE (EC 1.6.99.3) (Complex I-29KD) (CI-29KD) (Fragment).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatoxipha; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. Bintje; TISSUE=Tuber;
RX MEDLINE=94124587; PubMed=8294484;
RA Herz U., Schroeder W., Liddell A., Leaver C.J., Brennicke A.,
RA Grohmann L.;
RT "Purification of the NADH:ubiquinone oxidoreductase (complex I) of
the respiratory chain from the inner mitochondrial membrane of
Solanum tuberosum";
RL J. Biol. Chem. 269:2263-2269(1994).
CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
TO BE UBIQUINONE.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
CC -1- SUBCELLULAR LOCATION: MATRIX SIDE OF THE MITOCHONDRIAL INNER
MEMBRANE.
PIR; C49732; C49732.

KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2653 MW; 4B41B8FF83412F58 CRC64;

Query Match 21.2%; Score 21; DB 1; Length 23;
Best Local Similarity 40.0%; Pred. No. 2.7e+03;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFYG 16
Db 18 DYFYG 22

RESULT 14
SNBP_RAT STANDARD; PRT; 25 AA.
ID SNBP_RAT
AC P80968;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SM-11044 binding protein (Fragments).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC STRAIN=Wistar;
RX MEDLINE=97407910; PubMed=92611134;
RA Sugisawa T., Matsuzaki-Fujita M., Guillaume J.-L., Camoin L.,
RA Morooka S., Strosberg A.D.;
RT "Characterization of a novel iodocyanopindolol and SM-11044 binding
protein, which may mediate relaxation of depolarized rat colon
tonus";
RL J. Biol. Chem. 272:21244-21252(1997).
CC -1- FUNCTION: MAY MEDIATE RELAXATION OF DEPOLARIZED COLON TONUS.
CC -1- IT BINDS IODOCYANOPINDOLOL AND SM-11044.
CC -1- MISCELLANEOUS: THE ORDER OF THE FRAGMENTS IS UNKNOWN.
FT NON_TER 1 1
FT UNSURE 6 6 OR Y.
FT NON_CONS 18 19
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3177 MW; D14F0CB9B778C2CB CRC64;

Query Match 21.2%; Score 21; DB 1; Length 25;
Best Local Similarity 42.9%; Pred. No. 2.9e+03;
Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 12 EYFYGDF 18
Db 15 QYFYPX 21

RESULT 15
TKNB_RANRI STANDARD; PRT; 10 AA.
ID TKNB_RANRI
AC P29135;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurokinin A.
OS Rana ridibunda (Laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
OX NCBI_TaxID=8406;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=93075037; PubMed=1332683;
RA Wang Y., Badgery-Parker T., Lovas S., Chartrel N., Vaudry H.,
RA Burcher E., Conlon J.M.;
RT "Primary structure and receptor-binding properties of a neurokinin A-
related peptide from frog gut";

RL Biochem. J. 287:827-832(1992).
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
 CC MUSCLES.
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.

DR PIR; S27178; S27178.
 DR InterPro: IPR002040; Tachykinin.
 DR PROSITE: PS00267; TACHYKININ; 1.
 KW Tachykinin; Neuropeptide; Amidation; Amphibian skin.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1160 MW; 526B407059D5BAA7 CRC64;

Query Match 20.2%; Score 20; DB 1; Length 10;
 Best Local Similarity 25.0%; Pred. No. 1.8e+03;
 Matches 2; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 9 HQIEYFEG 16
 |:::|
 Db 1 HKLDSFIG 8

Search completed: April 23, 2003, 13:43:50
 Job time : 4.75169 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:21:27 ; Search time 27.9101 Seconds
(without alignments)
132.885 Million cell updates/sec

Title: US-09-836-073-14
Perfect score: 99
Sequence: 1 AALEAKICHQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rhodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	96.0	381	11	Q8CYP9 mus musculus
2	69	69.7	206	13	Q8QHI5
3	51	51.5	390	5	Q8TRV5
4	50	50.5	411	10	Q9RL36
5	50	50.5	422	10	Q94A38
6	49	49.5	391	2	P72382
7	49	49.5	391	2	P95709
8	49	49.5	391	16	Q99X57
9	48.5	49.0	928	10	Q9LJ02
10	48	48.5	165	5	O44678
11	48	48.5	545	10	O80567
12	47	47.5	119	10	Q92P89
13	47	47.5	396	5	O01806
14	46	46.5	568	16	Q9KKL7
15	44.5	44.9	839	5	Q18841
16	44	44.4	826	10	Q940X9

17	43.5	43.9	788	17	O58603
18	43	43.4	150	10	O80788
19	43	43.4	239	10	O9SMD8
20	43	43.4	381	16	O82632
21	43	43.4	389	10	O8S078
22	43	43.4	523	10	O94K80
23	43	43.4	596	5	Q8T6A8
24	43	43.4	654	4	Q96NN4
25	43	43.4	654	6	Q95KA5
26	43	43.4	658	5	Q18215
27	42.5	42.9	577	16	Q97DB6
28	42	42.4	96	10	Q9AUG1
29	42	42.4	337	2	Q8RK98
30	42	42.4	422	16	Q9X042
31	42	42.4	424	10	Q9S7W6
32	42	42.4	453	10	O8RWR2
33	42	42.4	505	17	O980D8
34	42	42.4	519	5	Q9VUI9
35	42	42.4	541	16	Q9ZKY5
36	42	42.4	542	16	O25534
37	42	42.4	658	10	Q9CAN3
38	42	42.4	671	4	Q9UL65
39	42	42.4	967	5	Q9B179
40	42	42.4	1064	5	Q18165
41	41	41.4	87	10	Q9SHW7
42	41	41.4	240	16	O8RE78
43	41	41.4	290	2	O68207
44	41	41.4	296	11	O9CTN3
45	41	41.4	343	2	Q9XBI7

ALIGNMENTS

RESULT 1

O9CYP9	PRELIMINARY;	PRT;	381 AA.
ID	O9CYP9		
AC	O9CYP9;		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DE	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Sjogren syndrome antigen B.		
GN	SSB.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=EMBRYO;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaado I., Pesole G., Quackenbush J.,		
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,		
RA	Hayashizaki Y.		
RT	Functional annotation of a full-length mouse cDNA collection.;		
RL	Nature 409:685-690(2001).		
DR	EMBL; AK017822; BAB30957.1; .		
DR	MGD; MGI:98423; Ssb.		

O58603	pyrococcus
O80788	arabidopsis
Q8SMD8	laminaria d
Q82632	salmonella
Q9S078	oryza sativ
Q94K80	arabidopsis
Q8T6A8	caenorhabdi
Q96NN4	homo sapien
Q95KA5	macaca fasc
Q18215	caenorhabdi
Q97DB6	clostridium
Q9AUG1	brassica ol
Q8RK98	mycoplasma
Q9X042	thermotoga
Q9S7W6	arabidopsis
Q8RWR2	arabidopsis
Q980D8	sulfolobus
Q9VUI9	drosofila
Q9ZKY5	helicobacte
O25534	helicobacte
Q9CAN3	arabidopsis
Q9B179	homo sapien
Q18165	caenorhabdi
Q9SHW7	arabidopsis
Q8RE78	fusobacteri
O68207	escherichia
Q9CTN3	mus musculu
Q9XBI7	bacillus ce

DR InterPro: IPR002344; Lupus Ia.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; rrm; 1.
 DR PRINTS: PRO0302; LUPUSLA.
 SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;

Query Match 96.0%; Score 95; DB 11; Length 381;
 Best Local Similarity 100.0%; Pred. No. 8.8e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALEAKICHQIEYYFGDF 18
 DB 12 ALEAKICHQIEYYFGDF 28
 |||||

RESULT 2
 Q8QH15 PRELIMINARY; PRT; 206 AA.
 AC Q8QH15;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Untranslated region binding-protein.
 GN UBP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA L'Ecuver T.J., Fang H.-L.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF467897; AAL76269.1; -;
 SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFE90E9 CRC64;

Query Match 69.7%; Score 69; DB 13; Length 206;
 Best Local Similarity 80.0%; Pred. No. 0.0011;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LEAKICHQIEYYFGD 17
 DB 13 LESKICQIEYYFGN 27
 |||||

RESULT 3
 Q8TV5 PRELIMINARY; PRT; 390 AA.
 AC Q8TV5;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE AT22034p.
 GN LA.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
 RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY075257; AAL68124.1; -;
 SQ SEQUENCE 390 AA; 44869 MW; 797FDE26B903C909 CRC64;

Query Match 51.5%; Score 51; DB 5; Length 390;
 Best Local Similarity 64.3%; Pred. No. 2.2;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 4 EAKICHQIEYYFGD 17
 DB 51 ERAIRQVEYYFGD 64
 |||||

RESULT 4
 Q9FL36 PRELIMINARY; PRT; 411 AA.
 ID Q9FL36;
 AC Q9FL36;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Similarity to RNA-binding protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA;
 RX MEDLINE-9834145; PubMed-9679202;
 RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
 RT features of the regions of 1,381,565 bp covered by twenty one
 RT physically assigned pl and TAC clones.";
 RL DNA Res. 5:131-145(1998).
 DR EMBL: AB010698; BAB11080.1; -;
 DR InterPro: IPR002344; Lupus_La.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; rrm; 1.
 DR PRINTS: PRO0302; LUPUSLA.
 DR SMART: SM00360; RRM; 1.
 DR PROSITE: PS0102; RRM; 1.
 SQ SEQUENCE 411 AA; 45655 MW; A2EF62EB5B9B099 CRC64;

Query Match 50.5%; Score 50; DB 10; Length 411;
 Best Local Similarity 60.0%; Pred. No. 3.3;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 LEAKICHQIEYYFGD 17
 DB 92 LNOKIIRQVEYYFSD 106
 |||||

RESULT 5
 Q9A38 PRELIMINARY; PRT; 422 AA.
 ID Q9A38;
 AC Q9A38;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE AT546250/WPL12.3
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Bhan J., Bowser L., Carinci P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
 RA Kirland-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miralinda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; AY050403; AAK91419.1; -
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; trm; 1.
DR PROSITE; PS50102; RRM; 1.
SQ SEQUENCE 422 AA; 46842 MW; 4EC4BBBF1E068F0E CRC64;

Query Match
Best Local Similarity 50.5%; Score 50; DB 10; Length 422;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 LEAKICHOIEYFVG 17
DB 103 LNKKIIRQVEYFSD 117

RESULT 6
P72382 PRELIMINARY; PRT; 391 AA.
AC P72382;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CAP8P.
GN CAP8P.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BECKER;
RX MEDLINE=96178981; PubMed=8606192;
RA Sau S., Sun J., Lee C.Y.;
RT "Cloning of type 8 capsule genes and analysis of gene clusters for the
RT production of different capsular polysaccharides in Staphylococcus
RT aureus.";
RL J. Bacteriol. 178:2118-2126(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BECKER;
RX MEDLINE=96178981; PubMed=8606192;
RA Sau S., Sun J., Lee C.Y.;
RT "Molecular characterization and transcriptional analysis of type 8
RT capsule genes in Staphylococcus aureus.";
RL J. Bacteriol. 179:1614-1621(1997).
DR EMBL; U73374; ABA49445.1; -.
DR HSSP; P27828; 1F6D.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMs; TIGR00236; wecB; 1.
SQ SEQUENCE 391 AA; 44247 MW; 1E8D9FAA9BC76F0D CRC64;

Query Match
Best Local Similarity 49.5%; Score 49; DB 2; Length 391;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIEYFVG 16
DB 366 RICEAIEYFVG 376

RESULT 7
P95709 PRELIMINARY; PRT; 391 AA.
AC P95709;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CAP5P.
GN CAP5P.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.

NCBI_TaxID=1280;
[1]
SEQUENCE FROM N.A.
STRAIN=REYNOLDS, AND NEWMAN;
MEDLINE=97388587; PubMed=9245821;
RA Sau S., Bhasin N., Wann E.R., Lee J.C., Foster T.J., Lee C.Y.;
RT "The Staphylococcus aureus allelic genetic loci for serotype 5 and 8
RT capsule expression contain the type-specific genes flanked by common
RL genes.";
RL Microbiology 143:0-0(0).
[2]
SEQUENCE FROM N.A.
STRAIN=REYNOLDS, AND NEWMAN;
MEDLINE=98101481; PubMed=9440531;
RA Kiser K.B., Lee J.C.;
RT "Staphylococcus aureus cap50 and cap5P genes functionally complement
RT mutations affecting enterobacterial common-antigen biosynthesis in
RT Escherichia coli.";
RL J. Bacteriol. 180:403-406(1998).
[3]
SEQUENCE FROM N.A.
STRAIN=REYNOLDS, AND NEWMAN;
MEDLINE=98125727; PubMed=9466251;
RA Bhasin N., Albus A., Michon F., Livolsi P.J., Park J.-S., Lee J.C.;
RT "Identification of a gene essential for O-acetylation of the
RT Staphylococcus aureus type 5 capsular polysaccharide.";
RL Mol. Microbiol. 27:9-21(1998).
[4]
SEQUENCE FROM N.A.
STRAIN=REYNOLDS, AND NEWMAN;
RA Bagga N., Wann E.R., Foster T.J., Lee J.C.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U91973; AAC46099.1; -.
DR HSSP; P27828; 1F6D.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMs; TIGR00236; wecB; 1.
SQ SEQUENCE 391 AA; 44328 MW; B0105F690DB7CF1D CRC64;

Query Match
Best Local Similarity 72.7%; Score 49; DB 2; Length 391;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIEYFVG 16
DB 366 RICEAIEYFVG 376

RESULT 8
Q99X57 PRELIMINARY; PRT; 391 AA.
AC Q99X57;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Capsular polysaccharide synthesis enzyme Cap5P.
GN CAP OR SAV0164 OR SA0159.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
SEQUENCE FROM N.A.
SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kato C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;

```

```

RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL: AP003358; BAB36326.1;
DR EMBL: AP003129; BAB41379.1;
DR HSSP: P27828; 1F6D.
DR InterPro: IPR003331; Epimerase_2.
DR Pfam: PF02350; Epimerase_2; 1.
DR TIGRFAMs: TIGR00236; wecB; 1.
KW Complete proteome.
SQ SEQUENCE 391 AA; 44372 MW; DDDF5FA715BCCEC CRC64;

Query Match 49.5%; Score 49; DB 16; Length 391;
Best Local Similarity 72.7%; Pred. No. 4.7;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIEYFYG 16
DB 366 RICEAIEYFYG 376

RESULT 9
Q9LJ02 PRELIMINARY; PRT; 928 AA.
AC Q9LJ02;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Similar to KIAA0731 protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone: P0499C11."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP001080; BAA90356.1;
SQ SEQUENCE 928 AA; 103745 MW; 53FC46E24A446EB4 CRC64;

Query Match 49.0%; Score 48.5; DB 10; Length 928;
Best Local Similarity 68.8%; Pred. No. 13;
Matches 11; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 3 LEAKICHOIEYF-GD 17
DB 278 LRKILTOVEYFSGD 293

RESULT 10
O44678 PRELIMINARY; PRT; 166 AA.
AC O44678;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 18.8 kDa protein.
GN Cl4C6.12.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
```

```

RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA David M., Wohldmann P., Bauer C., Antoniou B.;
RT "The sequence of C. elegans cosmid C14C6."
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF039051; AAB94258.1;
DR InterPro: IPR002542; DUF19.
DR Pfam: PF01579; DUF19; 1.
KW Hypothetical protein.
SQ SEQUENCE 166 AA; 18770 MW; 521379EFCB612487 CRC64;

Query Match 48.5%; Score 48; DB 5; Length 166;
Best Local Similarity 69.2%; Pred. No. 2.9;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 KICHOIEYFYGDF 18
DB 81 KICNVIEYMTGDF 93

RESULT 11
O80567 PRELIMINARY; PRT; 545 AA.
ID O80567; Q9C5X1;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Expressed protein (VirF-interacting protein FIP1) (Hypothetical 60.6
DE kDa protein) (At2g43970/F6E13.10).
GN AT2G43970 OR F6E13.10/At2G43970.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
RA Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Mayda E., Tzfira T., Citovsky V.;
RT "Arabidopsis thaliana VirF-interacting protein FIP1."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
```

Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,
 Theologis A.;
 "Full length cDNA of gene F6E13.10/At2g43970 (GI:3212854).";
 Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 [6]
 SEQUENCE FROM N.A.
 RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Shinn P.,
 Ban J., Bowser L., Carninci P., Chung M.K., Goldsmith A.D.,
 Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 Nguyen M., Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T.,
 Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 Yu G., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 "Arabidopsis cDNA clones";
 Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 [7]
 SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Ban J.,
 Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
 Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
 Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,
 Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
 Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,
 Davis R.W., Theologis A., Ecker J.R.;
 "Arabidopsis cDNA clones";
 Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 [8]
 RA EMBL; AC004005; AAC23405.2; -;
 DR EMBL; AF332565; AAK06847.1; -;
 DR EMBL; AY056238; AAL07087.1; -;
 DR EMBL; AF375410; AAK52994.1; -;
 DR EMBL; AF367277; AAK56266.1; -;
 DR InterPro; IPR002344; Lupus_La.
 DR PRINTS; PR00302; LUPUSLA.
 KW Hypothetical protein.
 SQ SEQUENCE 545 AA; 60589 MW; E1A933261FE1ED80 CRC64;
 Query Match 48.5%; Score 48; DB 10; Length 545;
 Best Local Similarity 66.7%; Pred. No. 9.5;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 6 KICHQIEYYFGD 17
 ||:|||||
 DB 196 KIVNQVEYFSD 207
 ||:|||||
 RESULT 12
 Q92P89
 ID Q92P89 PRELIMINARY; PRT; 119 AA.
 AC Q92P89
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Probable intron maturase (Maturase K) (Fragment).
 GN MATK.
 OS Neurotheca loeseloides.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Gentianales; Gentianaceae; Neurotheca.
 OX NCBI_TaxID=82724;
 [1]
 SEQUENCE FROM N.A.
 RA Thiv M., Kadereit J.W.;
 "The phylogenetic relationships and evolution of the Canarian laurel
 forest endemic Ixanthus viscosus (Ait.) Griseb. (Gentianaceae):
 evidence from matK and ITS sequences.";
 Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
 CC INTRONS (BY SIMILARITY).
 CC -1- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
 CC AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
 CC MITOCHONDRIAL INTRONS
 CC EMBL; AJ010524; CA837046.1; -;
 DR InterPro; IPR002866; MatK_N.

DR Pfam; PF01824; MatK_N; 1.
 KW Chloroplast; mRNA processing.
 FT NON_TER 1
 FT NON_TER 119 119
 SQ SEQUENCE 119 AA; 14383 MW; 18C631F7E7A5039D CRC64;
 Query Match 47.5%; Score 47; DB 10; Length 119;
 Best Local Similarity 50.0%; Pred. No. 3.1;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 3 LEAKICHQIEYYFGD 18
 ||:|||||
 DB 75 LENQLCHLRSYFFGDF 90
 ||:|||||
 RESULT 13
 O01806
 ID O01806 PRELIMINARY; PRT; 396 AA.
 AC O01806
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE C44E4.4 protein.
 GN C44E4.4
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE-99069613; PubMed-9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Sammons L., Wohldmann P., Gillam B.;
 "The sequence of C. elegans cosmid C44E4.";
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF003140; AAB54169.1; -;
 DR InterPro; IPR002344; Lupus_La.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
 SQ SEQUENCE 396 AA; 43631 MW; A2D828A4FAA3C34 CRC64;
 Query Match 47.5%; Score 47; DB 5; Length 396;
 Best Local Similarity 57.1%; Pred. No. 10;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 4 EAKICHQIEYYFGD 17
 ||:|||||
 DB 13 DOKIKQLEYFEGN 26
 ||:|||||
 RESULT 14
 Q9KKL7
 ID Q9KKL7 PRELIMINARY; PRT; 568 AA.
 AC Q9KKL7
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Response regulator.
GN VCA1086.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
DR EMBL; AE004434; AAF96979.1; -;
DR HSSP; P52934; IOMP.
DR TIGR; VCA1086;
DR InterPro; IPR001932; PP2C-like.
DR InterPro; IPR001789; Response_reg.
DR Pfam; PF00072; response_reg; 1.
DR ProDom; PD000039; Response_reg; 1.
DR SMART; SM00331; PP2C_SIG; 1.
DR SMART; SM00448; REC; 1.
KW Phosphorylation; Sensory transduction; Complete proteome.
SQ SEQUENCE 568 AA; 64331 MW; 87A02A350D986F51 CRC64;

Query Match 46.5%; Score 46; DB 16; Length 568;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 LEAKICHQIEYFG 16
I :|||:|
DB 156 LMEEMCHQVEHFG 169

RESULT 15
Q18841 PRELIMINARY; PRT; 839 AA.
AC Q18841;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE C54G10.2 protein.
GN C54G10.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Matthews L.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z75532; CAA99812.1; -;
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR000862; RFCdomain.
DR Pfam; PF00004; AAA; 1.
SQ SEQUENCE 839 AA; 93127 MW; FA38AC691F6FFB00 CRC64;

Query Match 44.9%; Score 44.5; DB 5; Length 839;

Best Local Similarity 57.9%; Pred. No. 57;
Matches 11; Conservative 1; Mismatches 2; Indels 5; Gaps 1;
QY 3 LEAKI-----CHOIEYFG 16
||||| ||||:|
DB 382 LEAKIGELSGSHQIEQFFG 400

Search completed: April 23, 2003, 13:32:52
Job time : 27.9101 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:42 ; Search time 4.95506 Seconds

(without alignments)
150.669 Million cell updates/sec

Title: US-09-836-073-14

Perfect score: 99

Sequence: 1 AALEAKICHQIEYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	100.0	404	1 LA_BOVIN	P10881 bos taurus
2	99	100.0	408	1 LA_HUMAN	P05455 homo sapien
3	99	100.0	415	1 LA_RAT	P38656 rattus norv
4	95	96.0	415	1 LA_MOUSE	P32067 mus musculus
5	70	70.7	427	1 LAB_XENLA	P28049 xenopus lae
6	66	66.7	428	1 LAA_XENLA	P28048 xenopus lae
7	55	55.6	383	1 LA_AEDAL	Q26457 aedes albop
8	51	51.5	390	1 LA_DROME	P40796 drosophila
9	46	46.5	298	1 LAHL_SCHPO	P87058 schizosacch
10	43	43.4	506	1 TCMQ_PETCR	Q43033 petroselinu
11	42	42.4	242	1 GLNQ_BACST	P27675 bacillus st
12	42	42.4	482	1 LBP_RABIT	P17454 oryctolagus
13	41	41.4	251	1 FOL2_MOUSE	Q05685 mus musculu
14	41	41.4	490	1 IFT4_MOUSE	Q14879 homo sapien
15	41	41.4	573	1 CATY_YEAST	P06115 saccharomyc
16	41	41.4	1164	1 RPO2_COMPX	P17474 cowpox viru
17	41	41.4	1164	1 RPO2_VACCV	P19798 vaccinia vi
18	41	41.4	1164	1 RPO2_VARV	P33811 variola vir
19	41	41.4	1220	1 DPOL_HSVB	P28858 equine herp
20	41	41.4	1221	1 V143_HSVAC	P24307 autographa
21	40	40.4	264	1 GRAK_HUMAN	P49863 homo sapien
22	40	40.4	466	1 SRO9_YEAST	P25567 saccharomyc
23	40	40.4	602	1 EXSA_BUCAI	P57530 buchnera ap
24	40	40.4	1107	1 MYIA_MOUSE	P46735 mus musculu
25	40	40.4	1136	1 MYIA_RAT	Q05096 rattus norv
26	39	39.4	482	1 IFT5_HUMAN	Q13325 homo sapien
27	39	39.4	488	1 RBL_OLIU	P14959 olisthodisc
28	39	39.4	812	1 PLMN_MOUSE	P20918 mus musculu
29	39	39.4	1048	1 AGOL_ARATH	O04379 arabidopsis
30	38	38.4	176	1 IPYR_HAEIN	P44529 haemophilus
31	38	38.4	354	1 ALKE_ARATH	Q9sa98 arabidopsis
32	38	38.4	391	1 CYB_PARTE	P15585 paramecium
33	38	38.4	488	1 RBL_ECTSI	P24313 ectocarpus

Query Match 100.0% Score 99; DB 1; Length 404;

ALIGNMENTS

RESULT 1

LA_BOVIN
ID LA_BOVIN STANDARD; PRT; 404 AA.
AC P10881;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
DE homology.
GN SSB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=89202037; PubMed=2468131;
RA Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences for RNA-binding";
RL Nucleic Acids Res. 17:2233-2244(1989).
CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S, AND 7-2 RNAs.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE C-TERMINAL PART OF THE PROTEIN.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X13698; CAA31986.1; ..
CC PIR; S03849; S03849.
CC InterPro; IPR002344; Lupus_La.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS0102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; 1.
CC RNA-binding; Nuclear protein; Phosphorylation.
CC DOMAIN 111 187 RNA-BINDING (RRM).
CC SEQUENCE 404 AA; 46534 MW; 4EE30B5C262AD6A1 CRC64;

P23651 pylaiella l
O84893 chlamydia t
P36640 salmonella
Q24087 drosophila
O9xgw1 arabidopsis
P87061 schizosacch
P97609 rattus norv
O43593 homo sapien
P46531 homo sapien
O31288 buchnera ap
P37891 oryza sativ
P11515 triticum ae

34 38 38.4 488 1 RBL_PYLLI
35 38 38.4 591 1 Y875_CHLTR
36 38 38.4 902 1 ATMA_SALTY
37 38 38.4 926 1 MET9_DROME
38 38 38.4 988 1 PINH_ARATH
39 38 38.4 1147 1 TEAL_SCHPO
40 38 38.4 1181 1 HAIR_RAT
41 38 38.4 1189 1 HAIR_HUMAN
42 38 38.4 2556 1 NTC1_HUMAN
43 37.5 37.9 153 1 IBP_BUCTS
44 37.5 37.9 500 1 CBP3_ORYSA
45 37.5 37.9 500 1 CBP3_WHEAT

Best Local Similarity 100.0%; Pred. No. 6.8e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
IIIIIIIIIIIIIIIIIIII
Db 11 AALEAKICHQIEYFGDF 28

RESULT 2

LA_HUMAN STANDARD; PRT; 408 AA.

AC P05455;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lupus La protein (Sjoren syndrome type B antigen (SS-B)) (La
DE ribonucleoprotein) (La autoantigen).
GN SSB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89202037; PubMed=2468131;
RA Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
RT sequences for RNA-binding.";
RL Nucleic Acids Res. 17:2233-2244(1989).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89053970; PubMed=3192525;
RA Chambers J.C., Kenan D., Martin B.J., Keene J.D.;
RT "Genomic structure and amino acid sequence domains of the human La
RT autoantigen.";
RL J. Biol. Chem. 263:18043-18051(1988).
[3]
RP SEQUENCE OF 54-408 FROM N.A.
RX MEDLINE=88199081; PubMed=2452201;
RA Sturgess A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
RA Coppel R.S.;
RT "Characteristics and epitope mapping of a cloned human autoantigen
RT La.";
RL J. Immunol. 140:3212-3218(1988).
[4]
RP SEQUENCE OF 54-97 FROM N.A.
RX MEDLINE=85166283; PubMed=3856888;
RA Chambers J.C., Keene J.D.;
RT "Isolation and analysis of cDNA clones expressing human lupus La
RT antigen.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119(1985).
[5]
RP FUNCTION.
RX MEDLINE=89251617; PubMed=2470590;
RA Gottlieb E., Steitz J.A.;
RT "Function of the mammalian La protein: evidence for its action in
RT transcription termination by RNA polymerase III.";
RL EMBO J. 8:851-861(1989).
[6]
RP PHOSPHORYLATION.
RX MEDLINE=97207017; PubMed=9054510;
RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Maraie R.J.;
RT "Phosphorylation of the human La antigen on serine 366 can regulate
RT recycling of RNA polymerase III transcription complexes.";
RL Cell 88:707-715(1997).
-1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
AND 7-2 RNAS.
-1- SUBCELLULAR LOCATION: Nuclear (Probable).
-1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE

C-TERMINAL PART OF THE PROTEIN.
-1- PTM: THE N-TERMINUS IS BLOCKED.
-1- DISEASE: SERA FROM PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS
OFTEN CONTAIN ANTIBODIES THAT REACT WITH THE NORMAL CELLULAR
LA PROTEIN AS IF THESE ANTIGEN WERE FOREIGN.
-1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC EMBL: X13697; CAA31985.1; -;
CC EMBL: J04205; AAS1885.1; -;
CC PIR: A31888; A31888.
CC PIR: A22956; A22956.
CC PIR: A31273; A31273.
CC PIR: S03848; S03848.
CC PIR: S11013; S11013.
CC Genew; HGNC:11316; SSB.
CC MIM; 109090; -;
CC InterPro: IPR002344; Lupus_La.
CC InterPro: IPR000504; RNA_rec_mot.
CC Pfam: PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS0102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; 1.
CC Systemic lupus erythematosus; RNA-binding; Phosphorylation;
KW Nuclear protein.
KW DOMAIN 111. 187 RNA-BINDING (RRM).
FT MOD_RES 366 366 PHOSPHORYLATION (BY CK2).
SQ SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;
Query Match 100.0%; Score 99; DB 1; Length 408;
Best Local Similarity 100.0%; Pred. No. 6.9e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
IIIIIIIIIIIIIIIIIIII
Db 11 AALEAKICHQIEYFGDF 28

RESULT 3

LA_RAT STANDARD; PRT; 415 AA.

AC P38636;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
DE homolog).
GN SSB OR SS-B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93246235; PubMed=7916708;
RA Samsel I., Troester H., Bartsch H., Schwemmler M., Igloi G.L.,
RA Bachmann M.;
RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La:
RT detection of species-specific variations.";
RL Gene 126:265-268(1993).
-1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
AND 7-2 RNAS.
-1- SUBCELLULAR LOCATION: Nuclear (Probable).
-1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE

Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 LEAKICHQIEYFGD 17
 ||| |::|::|::|
 Db 43 LEASTIRQIEYFGD 57

RESULT 8
 LA_DROME
 ID LA_DROME STANDARD; PRT; 390 AA.
 AC P40796; Q24375; Q9VIN2;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 GN LA OR CG10922.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Canton-S; TISSUP-Ovary;
 EX MEDLINE-94309632; PubMed-8035794;
 RA Bal C., Li Z., Iolles P.P.;
 RT "Developmental characterization of a Drosophila RNA-binding protein homologous to the human systemic lupus erythematosus-associated La/SS-B autoantigen.";
 RL Mol. Cell. Biol. 14:5123-5129(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94309661; PubMed-8035818;
 RA Yoo C.J., Wolin S.L.;
 RT "La proteins from Drosophila melanogaster and Saccharomyces cerevisiae: a yeast homolog of the La autoantigen is dispensable for growth.";
 RL Mol. Cell. Biol. 14:5412-5424(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Berkeley;
 RX MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Meyers A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Flossler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J., Hootin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclele J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTION TERMINATION BY RNA POLYMERASE III. BINDS RNA AND DNA. BINDS TO PRECURSORS OF RNA POLYMERASE III TRANSCRIPTS. MAY PLAY A SPECIALIZED ROLE DURING FLY DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT EMBRYONIC, LARVAL, PUPAL, AND ADULT DEVELOPMENT. EXPRESSION THROUGHOUT THE EMBRYO IS FOLLOWED BY A RESTRICTED PATTERN OF MESODERMAL EXPRESSION THAT IS LATER CONFINED TO THE VISCERAL MESODERM, GONADS, GUT, AND SALIVARY GLANDS.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 DR EMBL; U07652; AAA20518.1;
 DR EMBL; L32988; AAA21776.1;
 DR EMBL; AE003666; AAF53885.1;
 DR FlyBase; FBgn0011638; La.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR00504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 KW RNA-binding; Nuclear protein; DNA-binding.
 FT DOMAIN 149 234
 RNA-BINDING (RRM).
 A -> T (IN REF. 1).
 FT CONFLICT 169 169
 KH -> NS (IN REF. 1).
 FT CONFLICT 182 183
 A -> R (IN REF. 1).
 FT CONFLICT 283 283
 K -> N (IN REF. 1).
 FT CONFLICT 329 329
 SQ SEQUENCE 390 AA; 44884 MW; A809928B90446A5 CRC64;
 Query Match 51.5%; Score 51; DB 1; Length 390;
 Best Local Similarity 64.3%; Pred. NO. 0.51; Mismatches 4; Indels 0; Gaps 0;
 Matches 9; Conservative 1;
 Qy 4 EAKICHQIEYFGD 17
 ||| |::|::|::|
 Db 51 ERAIRQIEYFGD 64

RESULT 9
 LAHL SCHPO
 ID LAHL SCHPO STANDARD; PRT; 298 AA.
 AC P87058; Q10458; Q13362;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 GN SLAI OR SPAC57A10.10C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-98067398; PubMed-9404894;

```

Db      64 EAEVLKQVEYFSD 77

RESULT 10
TCMO_PETCR
ID TCMO_PETCR STANDARD; PRT; 506 AA.
AC Q43033;
AD
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trans-cinnamate 4-monooxygenase (EC 1.14.13.11) (Cinnamic acid
DE 4-hydroxylase) (CA4H) (C4H) (P450CA4H) (Cytochrome P450 73).
GN CYP73A10 OR CYP73.
OS Petroselinum crispum (Parsley) (Petroselinum hortense).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
ON NCBI_TaxID=4043;
RX [1];
RP SEQUENCE FROM N.A.
RA MEDLINE=95320184; PubMed=7597051;
RR Logemann E., Parniske M., Hahlbrock K.;
RT "Modes of expression and common structural features of the complete
RT phenylalanine ammonia-lyase gene family in parsley.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:5905-5909(1995).
CC -|- FUNCTION: CONTROLS CARBON FLUX TO PIGMENTS ESSENTIAL FOR
CC POLLINATION OR UV PROTECTION, TO NUMEROUS PHYTOALEXINS SYNTHESIZED
CC BY PLANTS WHEN CHALLENGED BY PATHOGENS, AND TO LIGNINS.
CC -|- CATALYTIC ACTIVITY: trans-cinnamate + NADPH + O(2) = 4-
CC hydroxycinnamate + NADP(+) + H(2)O.
CC -|- PATHWAY: Phenylpropanoid metabolism; second step.
CC -|- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@lsb-sib.ch).
-----
DR EMBL: L38898; AAC41660.1; --
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR DR POSITE; PS00086; CYTOCHROME_P450; 1.
DR Oxidoreductase; Monooxygenase; Heme; NADP.
FT BINDING 448 448 HEME (BY SIMILARITY).
SQ SEQUENCE 506 AA; 58047 MW; 32F00EE959D69CCF CRC64;

Query Match 43.4%; Score 43; DB 1; Length 506;
Best Local Similarity 42.9%; Pred.No.13;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 AKICHQIEYFGDF 18
... ||:|||||
Db 215 SRLAQSFHYFGDF 228

RESULT 11
GLNQ_BACST
ID GLNQ_BACST STANDARD; PRT; 242 AA.
AC P27675;
AD
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamine transport ATP-binding protein glnQ.
GN GLNQ.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
ON NCBI_TaxID=1422;
RX [1];
RP SEQUENCE FROM N.A.

```

STRAIN-NUB36;
MEDLINE-91310597; PubMed-1856180;
Wu L., Welker N.E.;
"Cloning and characterization of a glutamine transport operon of
Bacillus stearothermophilus NUB36: effect of temperature on
regulation of transcription.";1991).
J. Bacteriol. 173:4877-4888(1991).
-|- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
FOR GLUTAMINE. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
TRANSPORT SYSTEM.
-|- SUBCELLULAR LOCATION: Membrane-associated (Potential).
-|- INDUCTION: BY LACK OF GLUTAMINE.
-|- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M61017; AAA22483.1; --
PIR; A42478; A42478.
InterPro; IPR003593; AAA_ATPase.
InterPro; IPR003439; ABC_transportr.
Pfam; PF00005; ABC_tran; 1.
ProDom; PD000006; ABC_transportr; 1.
SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR Amino-acid transport; Transport; Membrane; ATP-binding.
FW NP_BIND 34 41 ATP (BY SIMILARITY).
KQ SEQUENCE 242 AA; 27436 MW; 102B1C5E332F31C8 CRC64;
SQ
Query Match 42.4%; Score 42; DB 1; Length 242;
Best Local Similarity 60.0%; Pred. No. 9.6;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 9 HQIEYVGFDF 18
||: ||: ||:
Db 5 HQVKNYGFDF 14

RESULT 12
LBP_RABIT
ID LBP_RABIT STANDARD; PRT; 482 AA.
AC PI7454;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Lipopolysaccharide-binding protein precursor (LBP).
GN LBP.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NC NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90385281; PubMed-2402637;
RA Schumann R.R., Leong S.R., Flagg G.W., Gray P.W., Wright S.D.,
RA Mathison J.C., Tobias P.S., Ulevitch R.J.;
RT "Structure and function of lipopolysaccharide binding protein.";
RL Science 249:1429-1431(1990).
RN [2]
RP SEQUENCE OF 27-66.
RP TISSUE=Serum;
RC MEDLINE-86306528; PubMed-2427635;
RX Tobias P.S., Soldau K., Ulevitch R.J.;
RT "Isolation of a lipopolysaccharide-binding acute phase reactant from
rabbit serum.";
RL J. Exp. Med. 164:777-793(1986).
CC -|- FUNCTION: BINDS TO THE LIPID A MOIETY OF BACTERIAL
CC LIPOPOLYSACCHARIDES (LPS), A GLYCOLIPID PRESENT IN THE OUTER
CC

```

CC MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. THE LBP/LPS COMPLEX SEEMS
CC TO INTERACT WITH THE CD14 RECEPTOR.
CC -!- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M35534; AAA99235.1; -.
CC DR PIR; B35843; B35843.
CC DR HSSP; P17213; 1BP1.
CC DR InterPro; IPR001124; LBP_BPI_CETP.
CC DR Pfam; PF01273; LBP_BPI_CETP; 1.
CC DR Pfam; PF02886; LBP_BPI_CETP_C; 1.
CC DR SMART; SM00328; BPI1; 1.
CC DR SMART; SM00329; BPI2; 1.
CC DR PROSITE; PS00400; LBP_BPI_CETP; 1.
CC KW Lipid transport; Antibiotic; Transmembrane; Glycoprotein; Signal.
CC FT SIGNAL 1 26
CC FT CHAIN 27 482 LIPOPOLYSACCHARIDE-BINDING PROTEIN.
CC FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CONFLICT 57 57 E > G (IN REF. 2).
CC FT CONFLICT 63 63 S > F (IN REF. 2).
CC FT CONFLICT 63 63 S > F (IN REF. 2).
CC SQ SEQUENCE 482 AA; 54001 MW; 628A6E0A647200C2 CRC64;

Query Match 42.4%; Score 42; DB 1; Length 482;
Best Local Similarity 80.0%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LEAKICHOIE 12
||:|||||
DB 194 LESKICRQIE 203

RESULT 13
FOL2_MOUSE
ID FOL2_MOUSE STANDARD; PRT; 251 AA.
AC Q05685;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Folate receptor beta precursor (FR-beta) (Folate receptor 2) (Folate-binding protein 2).
DE FOLR2 OR FOLBP2 OR FBP2.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9137339; PubMed=1894617;
RT Brigle K.E., Westin E.H., Houghton M.T., Goldman I.D.;
RA "Characterization of two cDNAs encoding folate-binding proteins from L1210 murine leukemia cells. Increased expression associated with a genomic rearrangement."
RT J. Biol. Chem. 266:17243-17249(1991).
RN [2]
RP SEQUENCE OF 1-49 FROM N.A.
RX MEDLINE=94140851; PubMed=8307991;
RT Brigle K.E., Seither R.L., Westin E.H., Goldman I.D.;
RA "Increased expression and genomic organization of a folate-binding protein homologous to the human placental isoform in L1210 murine leukemia cell lines with a defective reduced folate carrier."
RT J. Biol. Chem. 269:4267-4272(1994).
CC -!- FUNCTION: BINDS TO FOLATE AND REDUCED FOLIC ACID DERIVATIVES AND MEDIATES DELIVERY OF 5-METHYLTETRAHYDROFOLATE TO THE INTERIOR OF CELLS.

```

CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR
 CC (BY SIMILARITY).
 CC -1- PTM: EIGHT DISULFIDE BONDS ARE PRESENT (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE FOLATE RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M64817; AAA37599.1; -;
 CC EMBL: L25338; AAA37594.1; -;
 CC PIR: B40969; B40969.
 CC MGD: MGI:95569; F01r2.
 CC InterPro: IPR004269; Folate_rec.
 CC Pfam: PF03024; Folate_rec; 1.
 CC Receptor; Glycoprotein; Signal; Placenta; Folate-binding; Membrane;
 CC GPI-anchor; Multigene family.
 KW SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 227 FOLATE RECEPTOR BETA.
 FT PROPEP 228 251 REMOVED IN NATURE FORM (POTENTIAL).
 FT LIPID 227 227 GPI-ANCHOR (POTENTIAL).
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 251 AA; 28821 MW; 8404EACB1BFEC7 CRC64;

Query Match 41.4%; Score 41; DB 1; Length 251;
 Best Local Similarity 75.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 8 CHQIEYF 15
 || |||||
 Db 167 CHTFEYF 174

RESULT 14

IFT4_HUMAN STANDARD; PRT; 490 AA.
 AC Q14879; Q99634; Q9BSK7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Interferon-induced protein with tetratricopeptide repeats 4 (IFT4-4)
 DE (Interferon-induced 60 kDa protein) (IFI-60K) (ISG-60) (CIG49)
 DE (Retinoic acid-induced gene G protein) (RIG-G).
 GN IFT4 OR IFI60.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Foreskin;
 RX MEDLINE=98054347; PubMed=93911139;
 RA Zhu H., Cong J.P., Shenk T.;
 RT "Use of differential display analysis to assess the effect of human
 RT cytomegalovirus infection on the accumulation of cellular RNAs:
 RT induction of interferon-responsive RNAs".
 RL Proc. Natl. Acad. Sci. U.S.A. 94:13985-13990(1997).
 RN [2]
 RC SEQUENCE FROM N.A.
 RX MEDLINE=99047533; PubMed=9828129;
 RA de Veer M.J., Sim H., Whisstock J.C., Devenish R.J., Ralph S.J.;
 RT "IFI60/ISG60/IFT4, a new member of the human IFI54/IFT2 family of
 RT interferon-stimulated genes".
 RL Genomics 54:267-277(1998).
 RN [3]
 RC SEQUENCE FROM N.A.
 RA Yu M., Tong J., Mao M., Chen S., Chen Z.;
 RT "RIG-G, a novel gene induced by ATRA in acute promyelocytic

RT leukemia cells, is a new member of the ISG family.";
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RC SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE IFIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 8 TPR REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF026939; AAB95160.1; -;
 CC EMBL: AF083470; AAC63524.1; -;
 CC EMBL: U52513; AAB40606.1; -;
 CC EMBL: BC001383; AAH01383.1; -;
 CC EMBL: BC004977; AAH04977.1; -;
 CC Genew: HGNC:5411; IFIT4.
 CC MIM: 604650; -;
 CC InterPro: IPR001440; TPR.
 CC Pfam: PF00515; TPR; 5.
 CC SMART: SM0028; TPR; 3.
 KW Repeat; TPR repeat; Interferon induction.
 FT REPEAT 51 84 TPR 1.
 FT REPEAT 94 127 TPR 2.
 FT REPEAT 136 169 TPR 3.
 FT REPEAT 172 206 TPR 4.
 FT REPEAT 207 240 TPR 5.
 FT REPEAT 241 274 TPR 6.
 FT REPEAT 415 448 TPR 7.
 FT REPEAT 450 481 TPR 8.
 FT CONFLICT 44 44 F -> S (IN REF. 4; AAH04977).
 FT CONFLICT 359 359 Q -> Q (IN REF. 2).
 FT CONFLICT 435 435 MISSING (IN REF. 2).
 SQ SEQUENCE 490 AA; 55984 MW; B9F042B4DF7151D2 CRC64;
 Query Match 41.4%; Score 41; DB 1; Length 490;
 Best Local Similarity 43.8%; Pred. No. 28;
 Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 3 LEAKICHQIEYFEGDF 18
 || :|||: :|
 Db 34 LEDRVNCQIEFLNTEF 49

RESULT 15
 CATT_YEAST STANDARD; PRT; 573 AA.
 AC P06115;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Catalase T (EC 1.11.1.6).
 GN CTT1 OR YGR088W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RC SEQUENCE OF 12-573 FROM N.A.
 RX MEDLINE=87053966; PubMed=3536508;
 RA Hartig A., Ruis H.;
 RT "Nucleotide sequence of the Saccharomyces cerevisiae CTT1 gene and
 RT deduced amino-acid sequence of yeast catalase T".
 RL Eur. J. Biochem. 160:487-490(1986).
 RN [2]
 RC SEQUENCE FROM N.A.

Job time : 4.95506 secs

RA Wedler H., Scharfe M., Wedler E., Wambutt R.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Hernandez K., Weber N., Wipfli P., Schmidheini T.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-74 FROM N.A.
RX MEDLINE=86230135; PubMed=2423850;
RA Spevak W., Hartig A., Meindl P., Ruis H.;
RT "Heme control region of the catalase T gene of the yeast
RT Saccharomyces cerevisiae";
RL Mol. Gen. Genet. 203:73-78(1986).
RN [5]
RP SEQUENCE OF 405-409.
RC STRAIN=ATCC 44827 / SKQ2N;
RX MEDLINE=97190279; PubMed=9038161;
RA Norbeck J., Blomberg A.;
RT "Metabolic and regulatory changes associated with growth of
RT Saccharomyces cerevisiae in 1.4 M NaCl. Evidence for osmotic
RT induction of glycerol dissimilation via the dihydroxyacetone
RT pathway";
RL J. Biol. Chem. 272:5544-5554(1997).
RN [6]
RP SEQUENCE OF 405-409.
CC -|- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -|- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC -|- COFACTOR: HEME GROUP.
CC -|- SUBUNIT: HOMOTETRAMER.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- MISCELLANEOUS: THIS IS ONE OF TWO CATALASES IN S.CEREVISIAE; THE
CC OTHER IS CATALASE A, WHICH IS THE PEROXISOMAL FORM.
CC -|- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04625; CAA28298.1; -
DR EMBL; Z72873; CAA97090.1; -
DR EMBL; M30256; AAA34540.1; -
DR PIR; A26117; CSBYT.
DR HSP; P04040; 1F4J.
DR SGD; S0003320; CTT1.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Multigene family.
FT ACT_SITE 75 BY SIMILARITY.
FT ACT_SITE 148 BY SIMILARITY.
FT BINDING 362 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT CONFLICT 440 D -> V (IN REF. 1).
FT CONFLICT 550 C -> G (IN REF. 1).
SQ SEQUENCE 573 AA; 65741 MW; E03380543767377B CRC64;

Query Match 41.4%; Score 41; DB 1; Length 573;
Best Local Similarity 53.3%; Pred. No. 32;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 LEAKICHQIEYFGD 17
Db 191 LNPESIHQITYWFGD 205

Search completed: April 23, 2003, 13:28:07

GenCore version 5.1.4_p5_4578.
Copyright (c) 1993 - 2003 Compugen Ltd.

QM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:03 ; Search time 10.1124 Seconds
(without alignments)
171.119 Million cell updates/sec

Title: US-09-836-073-14

Perfect score: 99
Sequence: 1 AALEAKICHQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	100.0	404	1 S03849	ribonucleoprotein
2	99	100.0	408	1 A31888	ribonucleoprotein
3	99	100.0	415	1 JC1494	ribonucleoprotein
4	70	70.7	427	1 S33817	ribonucleoprotein
5	66	66.7	428	1 S33818	ribonucleoprotein
6	51	51.5	390	2 A53773	La/SS-B homolog D-
7	51	51.5	390	2 A53781	ribonucleoprotein
8	49	49.5	391	2 H89777	capsular polysacch
9	48	48.5	166	2 T32701	hypothetical prote
10	48	48.5	529	2 T00677	hypothetical prote
11	47	47.5	396	2 T30953	hypothetical prote
12	46	46.5	298	2 T38937	rna binding protei
13	46	46.5	298	2 T43542	RNA-binding protei
14	46	46.5	568	2 C82379	response regulato
15	44.5	44.9	839	2 T20230	hypothetical prote
16	43.5	43.9	788	2 A71076	hypothetical prote
17	43	43.4	150	2 A84488	hypothetical prote
18	43	43.4	381	2 AB0734	probable bacterio
19	43	43.4	506	2 T14907	trans-cinnamate 4-
20	43	43.4	658	2 T19487	hypothetical prote
21	42.5	42.9	577	2 D97337	mismatch repair pr
22	42	42.4	242	2 A42478	glutamine transpo
23	42	42.4	422	2 D72313	hypothetical prote
24	42	42.4	424	2 T46197	hypothetical prote
25	42	42.4	482	2 B35843	lipopolysaccharide
26	42	42.4	505	2 B90181	Na+/H+ antiporter
27	42	42.4	541	2 H71887	hypothetical prote
28	42	42.4	542	2 G64627	hypothetical prote
29	42	42.4	658	2 D96656	hypothetical prote

30 42 42.4 1131 2 T15617 hypothetical prote
31 41 41.4 87 2 C84494 hypothetical prote
32 41 41.4 251 2 B40969 folate-binding prote
33 41 41.4 419 2 AH0417 integrase [impor
34 41 41.4 469 2 C70357 hypothetical prote
35 41 41.4 506 2 F85016 probable RING zinc
36 41 41.4 573 1 CSBYT catalase (EC 1.11.
37 41 41.4 1156 2 T37411 RNA polymerase sub
38 41 41.4 1164 1 RNW28T DNA-directed RNA p
39 41 41.4 1164 1 RNW2CP DNA-directed RNA p
40 41 41.4 1164 2 T28566 A25R protein - var
41 41 41.4 1164 2 G72166 A24R protein - var
42 41 41.4 1164 2 G36850 DNA-directed DNA p
43 41 41.4 1220 1 DJBEC3 DNA-directed DNA p
44 41 41.4 1220 2 T42573 helicase (EC 3.6.1
45 41 41.4 1221 1 HJNVAV

ALIGNMENTS

RESULT 1

S03849
ribonucleoprotein La - bovine
N;Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C;Accession: S03849
R;Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A;Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequ
A;Reference number: S03848; MUID:89202037; PMID:2468131
A;Accession: S03849
A;Molecule type: mRNA
A;Residues: 1-404 <CHA>
A;Cross-references: EMBL:X13698; NID:g755; PIDN:CAA31986.1; PID:g756
A;Note: part of this sequence was confirmed by protein sequencing
C;Comment: This protein associates with a variety of small RNA molecules, most of
ay act as a transcription termination factor.
C;Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
C;Keywords: blocked amino end; phosphoprotein; RNA binding
F;112-178/Domain: ribonucleoprotein repeat homology <RRM>
F;113-118/Region: RNA-binding RNP2 motif
F;151-158/Region: RNA-binding RNPI motif
F;228-404/Domain: phosphorylated #status predicted <PHY>

Query Match 100.0%; Score 99; DB 1; Length 404;
Best Local Similarity 100.0%; Pred. No. 9.7e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
|||||
DB 11 AALEAKICHQIEYFGDF 28

RESULT 2

A31888
ribonucleoprotein La - human
N;Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome a
C;Species: Homo sapiens (man)
C;Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C;Accession: A31888; S03848; A2956; A61051; S11013; I55553; I70205; I70206; A3127
R;Chambers, J.C.; Kenan, D.; Martin, B.J.; Keene, J.D.
J. Biol. Chem. 263, 18043-18051, 1988

A;Title: Genomic structure and amino acid sequence domains of the human La autoant
A;Reference number: A31888; MUID:89053970; PMID:3192525
A;Accession: A31888

A;Molecule type: mRNA

A;Residues: 1-408 <CHA>

A;Cross-references: GB:J04205; NID:g178686; PIDN:AAA51885.1; PID:g178687

R;Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.

Nucleic Acids Res. 17, 2233-2244, 1989

A;Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequ

A:Reference number: S03848; MUID:89202037; PMID:2458131
 A:Accession: S03848
 A:Molecule type: mRNA
 A:Residues: 1-408 <CH2>
 A:Cross-references: EMBL:X13697; NID:G36414; PIDN:CAA31985.1; PID:G36415
 R:Chambers, J.C.; Keane, J.D.
 Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985
 A:Title: Isolation and analysis of cDNA clones expressing human lupus La antigen.
 A:Reference number: A22956; MUID:85166283; PMID:3856888
 A:Accession: A22956
 A:Molecule type: mRNA
 A:Residues: 45-97, 'LK' <CH3>
 A:Cross-references: GB:J04205
 A:Note: This sequence has been revised in reference A31888
 R:Nyman, U.; Ringertz, N.R.; Pettersson, I.
 Immunol. Lett. 22, 65-72, 1989
 A:Title: Demonstration of an amino terminal La epitope recognized by human anti-La sera.
 A:Reference number: A61051; MUID:89379261; PMID:2476379
 A:Accession: A61051
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-19, 'E', '21-47 <NYM>
 R:Sturgess, A.D.; Peterson, M.G.; McNeillage, L.J.; Whittingham, S.; Coppel, R.L.
 J. Immunol. 140, 3212-3218, 1988
 A:Title: Characteristics and epitope mapping of a cloned human autoantigen La.
 A:Reference number: S11013; MUID:88199081; PMID:2452201
 A:Accession: S11013
 A:Molecule type: mRNA
 A:Residues: 'E', '55-287, 'V', '289-408 <STU>
 A:Cross-references: EMBL:M20328; NID:G37456; PIDN:AAA36577.1; PID:G37457
 R:Kohsaka, H.; Yamamoto, K.; Fujii, H.; Miura, H.; Miyasaka, K.; Miyamoto, J.
 Clin. Invest. 85, 1566-1574, 1990
 A:Title: Fine epitope mapping the human SS-B/La protein: Identification of a distinct au
 A:Reference number: I55553; MUID:90237237; PMID:1692037
 A:Accession: I55553
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 81-107 <RES>
 A:Cross-references: GB:M35261; NID:G338491; PIDN:AAA36652.1; PID:G338495
 A:Accession: I70205
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 174-224 <RE2>
 A:Cross-references: GB:M35263; NID:G338492; PIDN:AAA36653.1; PID:G338496
 A:Accession: I70206
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 279-342 <RE3>
 A:Cross-references: GB:M35262; NID:G338493; PIDN:AAA36654.1; PID:G338497
 A:Comment: This protein associates with a variety of small RNA molecules, most of which
 ay act as a transcription termination factor.
 C:Genetics:
 A:Gene: GDB:SSB
 A:Cross-references: GDB:I25359; OMIM:109090
 A:Map position: 2
 A:Introns: 22/3; 151/3; 185/2; 209/2; 223/3; 264/3; 380/2
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:228-408/Domain: phosphorylated #status experimental <PHY>

Query Match 100.0%; Score 99; DB 1; Length 408;
 Best Local Similarity 100.0%; Pred. No. 9.8e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEYFGDF 18
 |||||
 Db 11 AALEAKICHQIEYFGDF 28

RESULT 3

JC1494
 ribonucleoprotein La - rat
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: JC1494; S25145
 R:Sensel, I.; Troester, H.; Bartsch, H.; Schwemmler, M.; Igloi, G.L.; Bachmann, M.
 Gene 126, 265-268, 1993
 A:Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection
 A:Reference number: JC1494; MUID:93246255; PMID:7916708
 A:Accession: JC1494
 A:Molecule type: mRNA
 A:Residues: 1-415 <SDM>
 A:Cross-references: GB:X67859; NID:G55778; PIDN:CAA48043.1; PID:G55779
 A:Experimental source: liver
 C:Comment: This protein associates with a variety of small RNA molecules, most of w
 ay act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:227-415/Domain: phosphorylated #status predicted <PHY>
 Query Match 100.0%; Score 99; DB 1; Length 415;
 Best Local Similarity 100.0%; Pred. No. 1e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEYFGDF 18
 |||||
 Db 11 AALEAKICHQIEYFGDF 28

RESULT 4
 S33817
 ribonucleoprotein La B - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33817; S28544
 R:Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression
 A:Reference number: S33817; MUID:93287095; PMID:8510143
 A:Accession: S33817
 A:Molecule type: mRNA
 A:Residues: 1-427 <SCH>
 A:Cross-references: EMBL:X68818; NID:G64875; PIDN:CAA48716.1; PID:G64876
 C:Comment: This protein associates with a variety of small RNA molecules, most of w
 ay act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:111-177/Domain: ribonucleoprotein repeat homology <RRM>
 F:112-117/Region: RNA-binding RNP2 motif
 F:150-157/Region: RNA-binding RNP1 motif
 F:227-427/Domain: phosphorylated #status predicted <PHY>
 Query Match 70.7%; Score 70; DB 1; Length 427;
 Best Local Similarity 80.0%; Pred. No. 0.00074;
 Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LEAKICHQIEYFGD 17
 | : ||| |||||
 Db 12 LDTKICEQIEYFGD 26

RESULT 5
 S33818
 ribonucleoprotein La A - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33818; S28545

R:Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.

J. Mol. Biol. 231, 196-204, 1993

A:Title: La proteins from *Xenopus laevis*. cDNA cloning and developmental expression.

A:Reference number: S33817; MUID:93287095; PMID:8510143

A:Accession: S33818

A:Molecule type: mRNA

A:Residues: 1-428 <SCH>

A:CROSS-references: EMBL:X68817; NID:G64873; PIDN:CAA48715.1; PID:G64874

C:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: phosphoprotein; RNA binding

F:112-178/Domain: ribonucleoprotein repeat homology <RRM>

F:113-118/Region: RNA-binding RNP2 motif

F:151-158/Region: RNA-binding RNP1 motif

F:228-428/Domain: phosphorylated #status predicted <PHY>

Query Match 66.7%; Score 66; DB 1; Length 428;

Best Local Similarity 78.6%; Pred. No. 0.0035;

Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17

DB 14 DKICEQIEYFGD 27

RESULT 6

A53773

La/SS-B homolog D-la - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000

C:Accession: A53773

R:Bei, C.; Li, Z.; Tollas, P.P.

Mol. Cell. Biol. 14, 5123-5129, 1994

A:Title: Developmental characterization of a *Drosophila* RNA-binding protein homologous to

A:Reference number: A53773; MUID:94309632; PMID:8035794

A:Accession: A53773

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-390 <BAI>

A:CROSS-references: GB:U07652; NID:G464019; PIDN:AAA20518.1; PID:G464020

C:Genetics:

A:Gene: FlyBase:La

A:CROSS-references: FlyBase:FBgn0011638

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: leucine zipper; RNA binding

Query Match 51.5%; Score 51; DB 2; Length 390;

Best Local Similarity 64.3%; Pred. No. 1;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17

DB 51 ERAIRQVEYFGD 64

RESULT 7

A53781

ribonucleoprotein La - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C:Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 20-Sep-1999

C:Accession: A53781

R:Yoo, C.J.; Wolin, S.L.

Mol. Cell. Biol. 14, 5412-5424, 1994

A:Title: La proteins from *Drosophila melanogaster* and *Saccharomyces cerevisiae*: a yeast

A:Reference number: A53781; MUID:94309661; PMID:8035818

A:Accession: A53781

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-390 <VOO>

A:CROSS-references: GB:L32988; NID:G488469; PID:G488470

C:Genetics:

A:Gene: FlyBase:La

A:CROSS-references: FlyBase:FBgn0011638

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: RNA binding

Query Match 51.5%; Score 51; DB 2; Length 390;

Best Local Similarity 64.3%; Pred. No. 1;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17

DB 51 ERAIRQVEYFGD 64

RESULT 8

H89777

capsular polysaccharide synthesis enzyme Cap5P [imported] - *Staphylococcus aureus* (

C:Species: *Staphylococcus aureus*

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: H89777

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.;

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu,

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: H89777

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-391 <KUR>

A:CROSS-references: GB:BA000018; PID:G13700080; PIDN:BA841379.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: Cap

C:Superfamily: lipopolysaccharide biosynthesis protein bplD

Query Match 49.5%; Score 49; DB 2; Length 391;

Best Local Similarity 72.7%; Pred. No. 2.2;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIEYFG 16

DB 366 RICEAIEYFG 376

RESULT 9

T32701

hypothetical protein C14C6.12 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T32701

R:David, M.; Wohldmann, P.; Bauer, C.; Antoniou, B.

submitted to the EMBL Data Library, December 1997

A:Description: The sequence of *C. elegans* cosmid C14C6.

A:Reference number: Z21210

A:Accession: T32701

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-166 <DAV>

A:CROSS-references: EMBL:AF039051; PIDN:AA894258.1; GSPDB:GN00023; CESP:C14C6.12

A:Experimental source: strain Bristol N2; clone C14C6

C:Genetics:

A:Gene: CESP:C14C6.12

A:Map position: 5

A:Introns: 42/1; 156/3

Query Match 48.5%; Score 48; DB 2; Length 166;

Best Local Similarity 69.2%; Pred. No. 1.4;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 KICHOIEYFGDF 18

DB 81 KICNVIEWMTGDF 93

C:Species: *Vibrio cholerae*
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: C82379
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, B.
 L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A:Reference number: AB2035; MUID:20406833; PMID:10952301
 A:Accession: C82379
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-568 <HEI>
 A:Cross-references: GB:AE004434; GB:AE003853; NID:g9658531; PIDN:AAF96979.1; GSPDB:GN001
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VCA1086
 A:Map position: 2

Query Match 46.5%; Score 46; DB 2; Length 568;
 Best Local Similarity 50.0%; Pred. No. 10;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 LEAKICHQIEYFPG 16

DB 156 LMEEMCHQVEHIFG 169

RESULT 15

T20230

hypothetical protein C54G10.2 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T20230

R:Matthews, L.

submitted to the EMBL Data Library, June 1996

A:Reference number: Z19240

A:Accession: T20230

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-839 <WIL>

A:Cross-references: EMBL:Z75532; PIDN:CAA99812.1; GSPDB:GN00023; CESP:C54G10.2

A:Experimental source: clone C54G10

C:Genetics:

A:Gene: CESP:C54G10.2

A:Map position: 5

A:Introns: 12/1; 34/3; 69/3; 326/1; 432/3; 535/3; 668/2

Query Match 44.9%; Score 44.5; DB 2; Length 839;
 Best Local Similarity 57.9%; Pred. No. 27;
 Matches 11; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

QY 3 LEAKI-----CHQIEYFPG 16

DB 382 LEAKIGELSGHQIEQFFG 400

Search completed: April 23, 2003, 13:34:37
 Job time: 11.1124 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:27:28 ; Search time 10.3146 Seconds
(without alignments)
139.834 Million cell updates/sec

Title: US-09-836-073-14
Perfect score: 99
Sequence: 1 AALEAKICHQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCF_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCFUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	100.0	18	9 US-09-836-073-1	Sequence 1, Appl
2	99	100.0	18	9 US-09-836-073-14	Sequence 14, Appl
3	99	100.0	460	9 US-10-102-806-695	Sequence 695, App
4	95	96.0	17	9 US-09-836-073-13	Sequence 13, Appl
5	93	93.9	18	9 US-09-836-073-9	Sequence 9, Appl
6	91	91.9	18	9 US-09-836-073-11	Sequence 11, Appl
7	91	91.9	18	9 US-09-836-073-12	Sequence 12, Appl
8	90	90.9	18	9 US-09-836-073-10	Sequence 10, Appl
9	88.5	89.4	18	9 US-09-836-073-16	Sequence 16, Appl
10	87	87.9	18	9 US-09-836-073-2	Sequence 2, Appl
11	87	87.9	18	9 US-09-836-073-3	Sequence 3, Appl
12	87	87.9	18	9 US-09-836-073-4	Sequence 4, Appl
13	84	84.8	18	9 US-09-836-073-7	Sequence 7, Appl
14	83	83.8	18	9 US-09-836-073-8	Sequence 8, Appl
15	76	76.8	18	9 US-09-836-073-15	Sequence 15, Appl
16	75	75.8	18	9 US-09-836-073-5	Sequence 5, Appl
17	68	68.7	18	9 US-09-836-073-6	Sequence 6, Appl
18	59	59.6	38	9 US-09-836-073-25	Sequence 25, Appl
19	59	59.6	38	9 US-09-836-073-25	Sequence 25, Appl

20	59	59.6	38	9 US-09-438-486-25	Sequence 25, Appl
21	59	59.6	38	9 US-10-053-758-25	Sequence 25, Appl
22	59	59.6	38	9 US-10-054-295-25	Sequence 25, Appl
23	59	59.6	38	9 US-10-054-611-25	Sequence 25, Appl
24	58.5	59.1	37	9 US-09-843-676-24	Sequence 24, Appl
25	58.5	59.1	37	9 US-09-766-253-24	Sequence 24, Appl
26	58.5	59.1	37	9 US-09-438-486-24	Sequence 24, Appl
27	58.5	59.1	37	9 US-10-053-758-24	Sequence 24, Appl
28	58.5	59.1	37	9 US-10-054-295-24	Sequence 24, Appl
29	58.5	59.1	37	9 US-10-054-611-24	Sequence 24, Appl
30	57	57.6	16	9 US-09-836-073-19	Sequence 19, Appl
31	48	48.5	39	9 US-09-843-676-26	Sequence 26, Appl
32	48	48.5	39	9 US-09-766-253-26	Sequence 26, Appl
33	48	48.5	39	9 US-09-438-486-26	Sequence 26, Appl
34	48	48.5	39	9 US-10-053-758-26	Sequence 26, Appl
35	48	48.5	39	9 US-10-054-295-26	Sequence 26, Appl
36	48	48.5	39	9 US-10-054-611-26	Sequence 26, Appl
37	44	44.4	18	9 US-09-836-073-17	Sequence 17, Appl
38	43	43.4	569	10 US-09-925-300-1583	Sequence 1583, Ap
39	42.5	42.9	18	9 US-09-836-073-18	Sequence 18, Appl
40	42	42.4	186	9 US-10-013-315-6	Sequence 6, Appl
41	42	42.4	456	10 US-09-861-400-6	Sequence 6, Appl
42	42	42.4	482	10 US-09-861-400-5	Sequence 5, Appl
43	41	41.4	33	9 US-09-776-724A-222	Sequence 222, App
44	40.5	40.9	2169	9 US-09-738-626-5455	Sequence 5455, Ap
45	40	40.4	204	9 US-10-102-806-743	Sequence 743, App

ALIGNMENTS

RESULT 1
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836, 073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316, 630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1
Query Match 100.0%; Score 99; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.1e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AALEAKICHQIEYFGDF 18
|||||
DB 1 AALEAKICHQIEYFGDF 18
RESULT 2
US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836, 073

```
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14
```

```
Query Match          100.0%; Score 99; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.1e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AALEAKICHQIEYFGDF 18
    |||||
Db 1 AALEAKICHQIEYFGDF 18
```

RESULT 3

```
; Sequence 895, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103PIC1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 695
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-806-695
```

```
Query Match          100.0%; Score 99; DB 9; Length 460;
Best Local Similarity 100.0%; Pred. No. 9.5e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AALEAKICHQIEYFGDF 18
    |||||
Db 63 AALEAKICHQIEYFGDF 80
```

RESULT 4

```
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Das, S.
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
```

US-09-836-073-13

```
Query Match          96.0%; Score 95; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 AALEAKICHQIEYFGDF 18
    |||||
Db 1 AALEAKICHQIEYFGDF 17
```

RESULT 5

```
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9
```

```
Query Match          93.9%; Score 93; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 AALEAKICHQIEYFGD 17
    |||||
Db 1 AALEAKICHQIEYFGD 17
```

RESULT 6

```
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11
```

```
Query Match          91.9%; Score 91; DB 9; Length 18;
Best Local Similarity 94.4%; Pred. No. 7.2e-08;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 AALEAKICHQIEYFGDF 18
    |||||
Db 1 AALEAKICHQIEYFGDF 18
```

RESULT 7

US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match 91.9%; Score 91; DB 9; Length 18;
Best Local Similarity 94.4%; Pred. No. 7.2e-08;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
|||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 8

US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match 90.9%; Score 90; DB 9; Length 18;
Best Local Similarity 94.4%; Pred. No. 1e-07;
Matches 17; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
|||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 9

US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match 89.4%; Score 88.5; DB 9; Length 19;
Best Local Similarity 94.7%; Pred. No. 1.9e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYFGDF 18
|||||
Db 1 AALEAKICHQIEYFGDF 19

RESULT 10

US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match 87.9%; Score 87; DB 9; Length 18;
Best Local Similarity 88.9%; Pred. No. 3e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18
|||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 11

US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match 87.9%; Score 87; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 3e-07;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHOIEYFGDF 18
|||:|||||:|||||
DB 1 AALQAKICHOIQYFGQF 18

RESULT 12

US-09-836-073-4

; Sequence 4, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-4

Query Match 87.9%; Score 87; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
|||||:|||||:|||||
DB 4 EAKICHOIEYFGDF 18

RESULT 13

US-09-836-073-7

; Sequence 7, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-7

Query Match 84.8%; Score 84; DB 9; Length 18;
Best Local Similarity 94.1%; Pred. No. 8.8e-07;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHOIEYFGD 17
|||||:|||||:|||||
DB 1 AALEAKICHOIEYQGD 17

RESULT 14

US-09-836-073-8

; Sequence 8, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-8

Query Match 83.8%; Score 83; DB 9; Length 18;
Best Local Similarity 88.9%; Pred. No. 1.3e-06;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALEAKICHOIEYFGDF 18
|||||:|||||:|||||
DB 1 AALEAKICHOIEQFGDF 18

RESULT 15

US-09-836-073-15

; Sequence 15, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 15

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Xenopus

US-09-836-073-15

Query Match 76.8%; Score 76; DB 9; Length 18;
Best Local Similarity 81.2%; Pred. No. 1.5e-05;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LEAKICHOIEYFGDF 18
||:|||||:|||||
DB 3 LDTKICEQIEYFGDF 18

Search completed: April 23, 2003, 13:38:20
Job time: 10.3146 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:24 ; Search time 11.2247 Seconds
(without alignments)
47.183 Million cell updates/sec

Title: US-09-836-073-14
Perfect score: 99
Sequence: 1 AALEAKICHQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	99	100.0	18	4	US-09-316-630-3
2	99	100.0	18	4	US-09-316-630-4
3	68	68.7	38	4	US-08-974-549A-214
4	59	59.6	38	3	US-08-851-843A-25
5	59	59.6	38	4	US-08-974-549A-215
6	59	59.6	38	4	US-08-854-050-25
7	59	59.6	38	4	US-08-430-323-25
8	58.5	59.1	37	3	US-08-851-843A-24
9	58.5	59.1	37	4	US-08-854-050-24
10	58.5	59.1	37	4	US-09-430-323-24
11	48	48.5	39	3	US-08-851-843A-26
12	48	48.5	39	4	US-08-974-549A-216
13	48	48.5	39	4	US-08-854-050-26
14	48	48.5	39	4	US-09-430-323-26
15	42	42.4	456	1	US-08-205-719-4
16	42	42.4	456	1	US-08-431-517F-6
17	42	42.4	482	4	US-08-431-517F-5
18	42	42.4	482	6	5245013-2
19	41	41.4	483	3	US-09-027-166-7
20	41	41.4	1220	2	US-08-680-326-38
21	40	40.4	431	1	US-08-311-023-2
22	39	39.4	754	2	US-08-941-262-1
23	39	39.4	755	2	US-08-941-262-3
24	39	39.4	811	4	US-09-199-637A-93
25	39	39.4	812	1	US-08-248-629A-1
26	39	39.4	812	1	US-08-451-932-1
27	39	39.4	812	1	US-08-452-260-1

28	39	39.4	812	1	US-08-326-785-1	Sequence 1, Appli
29	39	39.4	812	2	US-08-612-788-1	Sequence 1, Appli
30	39	39.4	812	2	US-08-605-598B-1	Sequence 1, Appli
31	39	39.4	812	2	US-08-429-743-1	Sequence 1, Appli
32	39	39.4	812	2	US-08-866-735-1	Sequence 1, Appli
33	39	39.4	812	3	US-09-066-028-1	Sequence 1, Appli
34	39	39.4	812	5	PCT-US95-05107-1	Sequence 1, Appli
35	38.5	38.9	740	4	US-09-323-872A-23	Sequence 23, Appli
36	38.5	38.9	864	4	US-09-323-872A-28	Sequence 28, Appli
37	38	38.4	984	4	US-09-287-354-2	Sequence 2, Appli
38	38	38.4	1068	1	US-08-537-210A-2	Sequence 2, Appli
39	38	38.4	1068	4	US-09-113-825-2	Sequence 3, Appli
40	38	38.4	1189	4	US-09-287-354-3	Sequence 3, Appli
41	38	38.4	1189	4	US-09-287-354-4	Sequence 4, Appli
42	38	38.4	1207	4	US-09-287-354-5	Sequence 5, Appli
43	38	38.4	2556	1	US-08-185-432-17	Sequence 17, Appli
44	38	38.4	2556	1	US-08-083-590A-20	Sequence 20, Appli
45	38	38.4	2556	3	US-08-532-384-20	Sequence 20, Appli

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match 100.0% Score 99; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.6e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AALEAKICHQIEYFGDF 18
|||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

;; PRIOR FILING DATE: 1997-10-06
;; PRIOR APPLICATION NUMBER: 08/321,427
;; PRIOR FILING DATE: 1994-10-11
;; PRIOR APPLICATION NUMBER: 60/086,527
;; PRIOR FILING DATE: 1998-05-22
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
;; OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 100.0%; Score 99; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.6e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHOIEYFGDF 18
|||||
DB 1 AALEAKICHOIEYFGDF 18

RESULT 3
US-08-974-549A-214
; Sequence 214, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 23-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/912,951
;; FILING DATE: 14-AUG-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/915,503
;; FILING DATE: 14-AUG-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US97/17618
;; FILING DATE: 01-OCT-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US97/17885
;; FILING DATE: 01-OCT-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph Ted
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002610US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 214:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-974-549A-214
Query Match 68.7%; Score 68; DB 4; Length 38;
Best Local Similarity 91.7%; Pred. No. 0.00014;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIEYFGDF 18
|||||
DB 1 ICHQIEYFGDF 12

RESULT 4
US-08-851-843A-25
; Sequence 25, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419

;; FILING DATE: 18-APR-1997
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-0029300S
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; US-08-851-843A-25

Query Match 59.6%; Score 59; DB 3; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0042; 1; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIEYFGD 17
|| |||||
Db 1 ICEQIEYFGD 11

RESULT 5
US-08-974-549A-215
; Sequence 215, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA: US 08/851,843
; FILING DATE: 06-MAY-1997

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/854,050
;; FILING DATE: 09-MAY-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/911,312
;; FILING DATE: 14-AUG-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/912,951
;; FILING DATE: 14-AUG-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/915,503
;; FILING DATE: 14-AUG-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US97/17618
;; FILING DATE: 01-OCT-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US97/17885
;; FILING DATE: 01-OCT-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph Ted
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-0026100S
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 215:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-974-549A-215

Query Match 59.6%; Score 59; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0042; 1; Indels 0; Gaps 0;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIEYFGD 17
|| |||||
Db 1 ICEQIEYFGD 11

RESULT 6
US-08-854-050-25
; Sequence 25, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536

;; PRIOR APPLICATION DATA: US 08/851,843
;; APPLICATION NUMBER: US 08/851,843
;; FILING DATE: 06-MAY-1997
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/846,017
;; FILING DATE: 25-APR-1997
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/844,419
;; FILING DATE: 18-APR-1997
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/844,419
;; FILING DATE: 18-APR-1997
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-0029300S
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; US-08-854-050-25

Query Match 59.6%; Score 59; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0042;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHOIEYFGD 17
Db 1 ICEQIEYFGD 11

RESULT 7
US-09-430-323-25
; Sequence 25, Application US/09430323
; Patent No. 6309667
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050

;; FILING DATE: 09-MAY-1997
;; APPLICATION NUMBER: US 08/851,843
;; FILING DATE: 06-MAY-1997
;; APPLICATION NUMBER: US 08/846,017
;; FILING DATE: 25-APR-1997
;; APPLICATION NUMBER: US 08/844,419
;; FILING DATE: 18-APR-1997
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-0029300S
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match 59.6%; Score 59; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0042;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHOIEYFGD 17
Db 1 ICEQIEYFGD 11

RESULT 8
US-08-851-843A-24
; Sequence 24, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002930US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 37 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; US-08-851-843A-24

Query Match 59.1%; Score 58.5; DB 3; Length 37;
Best Local Similarity 91.7%; Pred. No. 0.005;
Matches 11; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 7 ICHQIEYYFGDF 18
Db 1 ICHQ-EYFQDF 11

RESULT 9
US-08-854-050-24
; Sequence 24, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; LINGNER, Joachim
; NAKAMURA, Toru
; APPLICANT: NAKAMURA, Toru
; APPLICANT: CHAPMAN, Karen B.
; APPLICANT: MORIN, Gregg B.
; APPLICANT: HARLEY, Calvin
; APPLICANT: ANDREWS, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643

;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002930US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 37 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; US-08-854-050-24

Query Match 59.1%; Score 58.5; DB 4; Length 37;
Best Local Similarity 91.7%; Pred. No. 0.005;
Matches 11; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 7 ICHQIEYYFGDF 18
Db 1 ICHQ-EYFQDF 11

RESULT 10
US-09-430-323-24
; Sequence 24, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; LINGNER, Joachim
; NAKAMURA, Toru
; APPLICANT: NAKAMURA, Toru
; APPLICANT: CHAPMAN, Karen B.
; APPLICANT: MORIN, Gregg B.
; APPLICANT: HARLEY, Calvin
; APPLICANT: ANDREWS, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-OCT-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRAINEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-430-323-24

Query Match 59.1%; Score 58.5; DB 4; Length 37;
Best Local Similarity 91.7%; Pred. No. 0.005;
Matches 11; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 7 ICHQIEYFGDF 18
| | | | | | | | | |
Db 1 ICHQ-EYFYGDF 11

RESULT 11
US-08-851-843A-26
Sequence 26, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids

TYPE: amino acid
STRAINEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-26

Query Match 48.5%; Score 48; DB 3; Length 39;
Best Local Similarity 72.7%; Pred. No. 0.27;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICHQIEYFGD 17
| | | | | | | | | |
Db 1 ILRQVEYFGD 11

RESULT 12
US-08-974-549A-216
Sequence 216, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:

```
;; APPLICATION NUMBER: WO PCT/US97/17885
;; FILING DATE: 01-OCT-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002610US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 216:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 39 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-974-549A-216

Query Match 48.5%; Score 48; DB 4; Length 39;
Best Local Similarity 72.7%; Pred. No. 0.27;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIEYFGD 17
Db 1 ILRQVEYFGD 11

RESULT 13
US-08-854-050-26
; Sequence 26, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
```

```
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002930US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 26:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 39 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
US-08-854-050-26

Query Match 48.5%; Score 48; DB 4; Length 39;
Best Local Similarity 72.7%; Pred. No. 0.27;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIEYFGD 17
Db 1 ILRQVEYFGD 11

RESULT 14
US-09-430-323-26
; Sequence 26, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
```



```

; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 39 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: not relevant
;   TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-430-323-26

```

```

Query Match      48.5%; Score 48; DB 4; Length 39;
Best Local Similarity 72.7%; Pred. No. 0.27;
Matches      8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      7 ICHQIEYFGD 17
Db      1 ILQVEYFGD 11

```

```

RESULT 15
US-08-205-719-4
; Sequence 4, Application US/08205719
; Patent No. 5705398
; GENERAL INFORMATION:
; APPLICANT: Mintz, D. N.
; APPLICANT: Tobias, P. S.
; APPLICANT: Ulevitch, R. J.
; TITLE OF INVENTION: SYSTEM AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: THERAPEUTIC INHIBITORS OF LPS-MEDIATED SEPSIS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5705398th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/205,719
; FILING DATE: 02-MAR-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI201P; TSRI324.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 456 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-205-719-4

```

```

Query Match      42.4%; Score 42; DB 1; Length 456;
Best Local Similarity 80.0%; Pred. No. 36;
Matches      8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      3 LEAKICHOIE 12
Db      168 LESKICHOIE 177

```

Search completed: April 23, 2003, 13:36:32
Job time : 11.2247 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:23 ; Search time 28.5169 Seconds
(without alignments)
84.109 Million cell updates/sec

Title: US-09-836-073-14

Perfect score: 99
Sequence: 1 AALEAKICHQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /SID22/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	100.0	18	AA1980	Human la autoantigen
2	99	100.0	92	AA1981	Human secreted pro
3	99	100.0	408	AA1982	Human autoantigen
4	99	100.0	439	AA1983	Novel human diago
5	99	100.0	460	AA1984	Breast and ovarian
6	99	100.0	460	AA1985	Human ovarian anti
7	73	73.7	21	AA1986	La/SSB epitope 17.
8	51	51.5	390	AA1987	Drosophila melanog
9	48	48.5	913	AA1988	Arabidopsis thalia
10	48	48.5	923	AA1989	Arabidopsis thalia

11	48	48.5	993	21	AA1980	Arabidopsis thalia
12	44	44.4	135	22	AA1981	Novel human diago
13	44	44.4	224	23	AA1982	Human polypeptide
14	43	43.4	176	22	AA1983	Proionbacterium
15	43	43.4	544	22	AA1984	Human colon cancer
16	43	43.4	569	21	AA1985	Human prostate can
17	43	43.4	691	23	AA1986	Human secreted pro
18	43	43.4	697	22	AA1987	Human polypeptide
19	42	42.4	88	22	AA1988	DNA-cysteine methy
20	42	42.4	88	23	AA1989	Human diagnostic a
21	42	42.4	111	22	AA1990	Human MDDT SEQ ID
22	42	42.4	111	23	AA1991	H. pylori secreted
23	42	42.4	175	18	AA1992	Human polypeptide
24	42	42.4	191	22	AA1993	H. pylori GHP0 346
25	42	42.4	324	19	AA1994	Helicobacter polyp
26	42	42.4	324	19	AA1995	Arabidopsis thalia
27	42	42.4	395	21	AA1996	Arabidopsis thalia
28	42	42.4	395	21	AA1997	Arabidopsis thalia
29	42	42.4	415	21	AA1998	Arabidopsis thalia
30	42	42.4	415	21	AA1999	Arabidopsis thalia
31	42	42.4	424	21	AA2000	Arabidopsis thalia
32	42	42.4	424	21	AA2001	Arabidopsis thalia
33	42	42.4	456	19	AA2002	Rabbit LPS-binding
34	42	42.4	482	14	AA2003	Lapine gram-negati
35	42	42.4	482	15	AA2004	Lapine polysacchar
36	42	42.4	482	19	AA2005	Rabbit lipopolysac
37	42	42.4	482	20	AA2006	Rabbit lipopolysac
38	42	42.4	517	23	AA2007	Oestrogen-regulate
39	42	42.4	519	22	AA2008	Drosophila melanog
40	42	42.4	519	23	AA2009	Oestrogen-regulate
41	42	42.4	542	22	AA2010	H. pylori HPS065 P
42	42	42.4	557	22	AA2011	H. pylori HPS065 P
43	42	42.4	609	20	AA2012	An alternatively s
44	42	42.4	609	21	AA2013	Human Ras signalli
45	42	42.4	728	20	AA2014	A human MCG7 prote

ALIGNMENTS

RESULT 1
AA1980
ID AA1980 standard; peptide; 18 AA.

AC AA1980;

XX AA1980;

DT 14-MAR-2000 (first entry)

DE Human la autoantigen peptide (LAP).

XX La autoantigen; LAP; internal ribosome entry site; IRES; translation;

KW viral replication; IRNA; antiviral agent; picornavirus; flavivirus;

KW coronavirus; hepatitis virus; rhinovirus; adenovirus; coxsackie virus;

KW parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;

KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;

XX vesicular stomatitis virus.

OS Homo sapiens.

XX WO9961613-A2.

PD 02-DEC-1999.

XX 21-MAY-1999; 99WO-US11281.

XX 22-MAY-1998; 98US-0086527.

XX (REGC) UNIV CALIFORNIA.

XX Das S, Dasgupta A;

XX WPI; 2000-062712/05.

PT New yeast inhibitory peptide useful for inhibiting viral protein
 PT translation and replication -
 PS Claim 5; Page 57; 81pp; English.
 XX
 CC This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (IRNA see RAZ45200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rhinovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.
 XX

SQ Sequence 18 AA;
 Query Match 100.0%; Score 99; DB 21; Length 18;
 Best Local Similarity 100.0%; Pred. No. 5.3e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYYFGDF 18
 |||||
 DB 1 AALEAKICHQIEYYFGDF 18

RESULT 2
 AAG01351
 ID AAG01351 standard; Protein; 92 AA.
 AC AAG01351;
 XX
 DT 06-OCT-2000 (first entry)
 DE Human secreted protein, SEQ ID NO: 5432.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 DR N-PSDB; AAC01357.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13; SEQ ID 5432; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30

CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX

SQ Sequence 92 AA;
 Query Match 100.0%; Score 99; DB 21; Length 92;
 Best Local Similarity 100.0%; Pred. No. 3e-08;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYYFGDF 18
 |||||
 DB 11 AALEAKICHQIEYYFGDF 28

RESULT 3
 AAW03716
 ID AAW03716 standard; protein; 408 AA.
 XX
 AC AAW03716;
 XX
 DT 12-MAR-1997 (first entry)
 DE Human autoantigen La(SS-B).
 XX
 KW Autoimmune disease; La autoantigen; Sjogren's syndrome;
 KW systemic lupus erythematosus; diagnosis.
 XX
 OS Homo sapiens.
 XX
 PN US5541291-A.
 XX
 PD 30-JUL-1996.
 XX
 PF 31-DEC-1984; 84US-0687908.
 XX
 PR 27-MAY-1987; 87US-0054871.
 PR 31-DEC-1984; 84US-0687908.
 XX
 PA (UYDU-) UNIV DUKE.
 XX
 PI Keene JD;
 XX
 DR WPI; 1996-362015/36.
 XX
 PT Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
 PT overlap syndrome - useful for diagnosis and treatment of autoimmune
 PT diseases
 XX
 PS Disclosure; Columns 15-16; 21pp; English.
 XX
 CC The human lupus antigen (La) is diagnostic for Sjogren's syndrome,
 CC as well as occurring in systemic lupus erythematosus patients. The
 CC La protein is clinically related to the Ro protein that is highly
 CC common among autoimmune patients. La and Ro antigens sometimes
 CC reside on the same cellular ribonucleoprotein particle; most La
 CC patients contain some Ro antibodies and vice versa. La cDNA has
 CC been isolated from a human liver library.
 XX

SQ Sequence 408 AA;
 Query Match 100.0%; Score 99; DB 17; Length 408;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AALEAKICHQIEYYFGDF 18

Db 11 AALEAKICHQIEYFGDF 28
|||||

RESULT 4
ABG08417
ID ABG08417 standard; Protein; 439 AA.
AC ABG08417;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #8408.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
PI WPI; 2001-639362/73.
XX
DR N-PSDB; AAS72604.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID NO 38776; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic amino acid sequences of the invention.
Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 439 AA;
Query Match 100.0%; Score 99; DB 22; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.0e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEYFGDF 18
|||||
Db 41 AALEAKICHQIEYFGDF 58

RESULT 5
AAB58987
ID AAB58987 standard; Protein; 460 AA.
XX
AC AAB58987;
XX
DT 27-MAR-2001 (first entry)
XX
DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 695.
XX
KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor; antiparasitic; anticonvulsant; antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis; cardiovascular disorder; wound healing; neurological disease.
XX
OS Homo sapiens.
XX
PN WO200055173-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05881.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
PI WPI; 2000-611515/58.
XX
DR N-PSDB; AAF21890.
XX
PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases
XX
PS Claim 11; Page 1149-1150; 1299pp; English.
XX
CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antitumor; antiparasitic; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiac activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.

Sequence 460 AA;
Query Match 100.0%; Score 99; DB 21; Length 460;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEYFGDF 18
|||||

DB 63 AALEAKICHQIEYYFGDF 80

RESULT 6
ABP41511

ID ABP41511 standard; Protein; 460 AA.

XX AC ABP41511;

XX AC 22-AUG-2002 (first entry)

DT DT

XX DE Human ovarian antigen HVAF56, SEQ ID NO:2643.

XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

XX KW ovarian cancer; breast cancer; tumour; reproductive system disorder;

XX KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

XX KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;

XX KW inflammatory condition; immune disorder; blood disorder;

XX KW cardiovascular disorder; respiratory disorder; neurological disorder;

XX KW gastrointestinal disorder; urinary system disorder; drug screening;

XX KW gene therapy; chromosome mapping; forensic analysis;

XX KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;

XX KW antiinflammatory; gynaecological; reproductive.

XX OS Homo sapiens.

XX PN WO200200677-A1.

XX PD 03-JAN-2002.

XX PF 07-JUN-2001; 2001WO-US18569.

XX PR 07-JUN-2000; 2000US-209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX DR WPI; 2002-147878/19.

XX DR N-PSDB; ABQ54588.

XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,

XX PT useful in the prevention, treatment and diagnosis of cancer (e.g.

XX PT ovarian cancer), immune disorders, cardiovascular disorders and

XX PT neurological diseases.

XX PS Claim 11; SEQ ID NO 2643; 2922pp; English.

XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
encompasses polypeptides 90% identical and polynucleotides 95% identical
to the sequences of the invention. The invention additionally relates to
recombinant vectors and host cells comprising human ovarian antigen
polynucleotides, antibodies against human ovarian antigens, and the use
of ovarian antigen polynucleotides and polypeptides in diagnosing,
treating, prognosing or preventing various ovary and/or breast-related
disorders. Such conditions include ovarian cancer and breast cancer, and
metastatic tumours of ovarian or breast origin, reproductive system
disorders (e.g., infertility, disorders of pregnancy, anovulation,
polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
vaginitis), immune disorders (e.g., congenital and acquired
immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
blood-related disorders (e.g., anaemia), cardiovascular disorders,
respiratory disorders, neurological disorders, gastrointestinal disorders
and urinary system disorders. Ovarian antigen polypeptides and
polynucleotides may also be used in screening for compounds which
modulate ovarian antigen expression or activity. The polynucleotides may
further be used for gene therapy, chromosome mapping, in the
identification of individuals and in forensic analysis, and the
polypeptides may be used as food additives or to prepare antibodies
useful in disease diagnosis, drug targeting and phenotyping. The present

CC sequence represents a human ovarian antigen of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 460 AA;

Query Match 100.0%; Score 99; DB 23; Length 460;
Best Local Similarity 100.0%; Pred. NO. 1.7e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYYFGDF 18
|||||
DB 63 AALEAKICHQIEYYFGDF 80

RESULT 7
AAR43394

ID AAR43394 standard; peptide; 21 AA.

XX AC AAR43394;

XX DT 12-MAY-1994 (first entry)

XX DE La/SSB epitope 17.

XX KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
KW nuclear ribonucleoprotein; rRNP; Sm B/B'; polypeptide; antigen; D;
KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

XX OS Homo sapiens.

XX PN WO9321223-A.

XX PD 28-OCT-1993.

XX PF 13-APR-1993; 93WO-US03484.

XX PR 13-APR-1992; 92US-0867819.

XX PA (OKLA) UNIV OKLAHOMA STATE.

XX PI Harley JB;

XX DR WPI; 1993-351658/44.

XX PT New linear epitope(s) for human auto-antibodies - from the
PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
PT for diagnosing and treating auto-immune disorders e.g. systemic
PT lupus erythematosus

XX CC Claim 1; Page 30; 43pp; English.

XX CC The sequences given in AAR43391-562 are linear epitopes which are
CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
CC the 70 kD nuclear ribonucleoprotein (rRNP) and the Sm B/B'
CC polypeptide. These antigens are common in systemic lupus
CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
CC of proteins has been shown to have several molecular forms which are
CC defined by the molecular weight of the antigen identified. The major
CC form has a molecular weight of 60 kD and two additional forms have
CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
CC group of autoantibodies and binds small RNAs with a polynucleotide
CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
CC phosphoprotein which associates with RNA polymerase III transcripts.
CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
CC U5 RNA. Anti-Sm antibodies may be directed against one or a
CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
CC used for preventing, treating or screening autoimmune disorders,
CC especially SLE or Sjogrens syndrome (SS). They bind to a human

PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 48.5%; Score 48; DB 21; Length 913;
 Best Local Similarity 66.7%; Pred. No. 45;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIEYFEGD 17
 || : ||||| |
 Db 116 KIVNQVEYFSD 127

RESULT 10
 AAG47713
 ID AAG47713 standard; Protein; 923 AA.
 XX
 AC AAG47713;
 XX
 DT 18-OCT-2000 (first entry)

XX	Arabidopsis thaliana protein fragment SEQ ID NO: 50167.	PR	21-JUN-1999;	99US-0139817.
DE		PR	22-JUN-1999;	99US-0139899.
XX		PR	23-JUN-1999;	99US-0140353.
KW	Protein identification; signal transduction pathway; metabolic pathway;	PR	23-JUN-1999;	99US-0140354.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	PR	24-JUN-1999;	99US-0140695.
KW	termination sequence.	PR	28-JUN-1999;	99US-0140823.
XX		PR	29-JUN-1999;	99US-0140991.
OS	Arabidopsis thaliana.	PR	30-JUN-1999;	99US-0141287.
XX		PR	01-JUL-1999;	99US-0141842.
PN	EP1033405-A2.	PR	01-JUL-1999;	99US-0142154.
XX		PR	02-JUL-1999;	99US-0142055.
PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
XX		PR	09-JUL-1999;	99US-0142920.
PF	25-FEB-2000; 2000EP-0301439.	PR	12-JUL-1999;	99US-0142977.
XX		PR	13-JUL-1999;	99US-0143542.
XX	25-FEB-1999; 99US-0121825.	PR	14-JUL-1999;	99US-0143624.
PR	05-MAR-1999; 99US-0123180.	PR	15-JUL-1999;	99US-0144005.
PR	09-MAR-1999; 99US-0123548.	PR	16-JUL-1999;	99US-0144085.
PR	23-MAR-1999; 99US-0125788.	PR	16-JUL-1999;	99US-0144086.
PR	23-MAR-1999; 99US-0126264.	PR	19-JUL-1999;	99US-0144325.
PR	29-MAR-1999; 99US-0126785.	PR	19-JUL-1999;	99US-0144331.
PR	01-APR-1999; 99US-0127462.	PR	19-JUL-1999;	99US-0144332.
PR	06-APR-1999; 99US-0128234.	PR	19-JUL-1999;	99US-0144333.
PR	08-APR-1999; 99US-0128714.	PR	19-JUL-1999;	99US-0144334.
PR	16-APR-1999; 99US-0129845.	PR	19-JUL-1999;	99US-0144335.
PR	19-APR-1999; 99US-0130077.	PR	20-JUL-1999;	99US-0144352.
PR	21-APR-1999; 99US-0130449.	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999; 99US-0130510.	PR	20-JUL-1999;	99US-0144884.
PR	28-APR-1999; 99US-0130891.	PR	21-JUL-1999;	99US-0144884.
PR	28-APR-1999; 99US-0131449.	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999; 99US-0132048.	PR	21-JUL-1999;	99US-0145088.
PR	30-APR-1999; 99US-0132407.	PR	22-JUL-1999;	99US-0145085.
PR	04-MAY-1999; 99US-0132484.	PR	22-JUL-1999;	99US-0145087.
PR	05-MAY-1999; 99US-0132485.	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999; 99US-0132486.	PR	22-JUL-1999;	99US-0145192.
PR	06-MAY-1999; 99US-0132487.	PR	23-JUL-1999;	99US-0145145.
PR	07-MAY-1999; 99US-0132863.	PR	23-JUL-1999;	99US-0145218.
PR	11-MAY-1999; 99US-0134256.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999; 99US-0134218.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999; 99US-0134219.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999; 99US-0134221.	PR	27-JUL-1999;	99US-0145918.
PR	14-MAY-1999; 99US-0134370.	PR	27-JUL-1999;	99US-0145919.
PR	18-MAY-1999; 99US-0134768.	PR	28-JUL-1999;	99US-0145951.
PR	19-MAY-1999; 99US-0134941.	PR	02-AUG-1999;	99US-0146386.
PR	20-MAY-1999; 99US-0135124.	PR	02-AUG-1999;	99US-0146388.
PR	21-MAY-1999; 99US-0135353.	PR	02-AUG-1999;	99US-0146389.
PR	24-MAY-1999; 99US-0135629.	PR	03-AUG-1999;	99US-0147038.
PR	25-MAY-1999; 99US-0136021.	PR	04-AUG-1999;	99US-0147204.
PR	27-MAY-1999; 99US-0136392.	PR	04-AUG-1999;	99US-0147302.
PR	28-MAY-1999; 99US-0136782.	PR	05-AUG-1999;	99US-0147192.
PR	01-JUN-1999; 99US-0137222.	PR	05-AUG-1999;	99US-0147260.
PR	03-JUN-1999; 99US-0137528.	PR	06-AUG-1999;	99US-0147303.
PR	04-JUN-1999; 99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999; 99US-0137724.	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999; 99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999; 99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999; 99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999; 99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999; 99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999; 99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999; 99US-0139452.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999; 99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999; 99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999; 99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999; 99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999; 99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999; 99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999; 99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999; 99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999; 99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999; 99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999; 99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999; 99US-0139763.	PR	27-AUG-1999;	99US-0151066.


```
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147433.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 48.5%; Score 48; DB 21; Length 993;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 6 KICHOIEYFSD 17
Db 196 KIVNOVEYFSD 207
||:|||||
||:|||||

RESULT 12
ABG27058
ID ABG27058 standard; Protein; 135 AA.
XX
AC ABG27058;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #27049.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
```

```

PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
PA Dmanac RT, Liu C, Tang YT;
PI WPI: 2001-639362/73.
DR N-PSDB; AAS91245.
DR
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 20; SEQ ID NO 57417; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 135 AA;
Query Match 44.4%; Score 44; DB 22; Length 135;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 4 EAKICHQIEY 13
Db 9 EKRICHQIEF 18
RESULT 13
ABB89645
ID ABB89645 standard; Protein; 224 AA.
XX
AC ABB89645;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 2021.
XX
KW Cystostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antinflammatory; antiulcer;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US16450.

XX 19-MAY-2000; 2000US-205515P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Birse CE, Rosen CA;
XX WPI: 2002-122018/16.
XX N-PSDB; ABL90054.
XX
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders
XX
PS Claim 11; SEQ ID NO 2021; 208ipp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 224 AA;
Query Match 44.4%; Score 44; DB 23; Length 224;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 4 EAKICHQIEY 13
Db 112 EKRICHQIEF 121
RESULT 14
AAU52458
ID AAU52458 standard; Protein; 176 AA.
XX
AC AAU52458;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #13354.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.

```

XX PA (CORI-) CORIXA CORP.

XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX DR WPI; 2001-616774/71.

XX DR N-PSDB; AAS59555.

XX PT Propionibacterium acnes polypeptides and nucleic acids useful for

XX PT vaccinating against and diagnosing infections, especially useful for

XX PT treating acne vulgaris -

XX PS Example 1; SEQ ID No 13653; 1069pp; English.

XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

XX CC polypeptides. The proteins and their associated DNA sequences are used in

XX CC the treatment, prevention and diagnosis of medical conditions caused by

XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

XX CC pustulosis, hypertostosis and osteomyelitis), uveitis and endophthalmitis.

XX CC P. acnes is also involved in infections of bone, joints and the central

XX CC nervous system, however it is particularly involved in the inflammatory

XX CC lesions associated with acne vulgaris. A method for detecting the

XX CC presence or absence of P. acnes in a patient comprises contacting a

XX CC sample with a binding agent that binds to the proteins of the invention

XX CC and determining the amount of bound protein in the sample. The

XX CC polypeptides may be used as antigens in the production of antibodies

XX CC specific for P. acnes proteins. These antibodies can be used to

XX CC downregulate expression and activity of P. acnes polypeptides and

XX CC therefore treat P. acnes infections. The antibodies may also be used as

XX CC diagnostic agents for determining P. acnes presence, for example, by

XX CC enzyme linked immunosorbent assay (ELISA).

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 176 AA;

Query Match 43.4%; Score 43; DB 22; Length 176;

Best Local Similarity 58.3%; Pred. No. 48;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 LEAKICHQIEY 14

Db 66 IEPRICHNIEAY 77

RESULT 15

AAG75090

XX AC AAG75090 standard; Protein; 544 AA.

XX AC AAG75090;

XX DT 03-SEP-2001 (first entry)

XX DE Human colon cancer antigen protein SEQ ID NO:5854.

XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX KW colorectal carcinoma.

XX OS Homo sapiens.

XX PN WO200122920-A2.

XX PD 05-APR-2001.

XX PF 28-SEP-2000; 2000WO-US26524.

XX PR 29-SEP-1999; 99US-0157137.

XX PR 03-NOV-1999; 99US-0163280.

XX PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

DR N-PSDB; AAH34495.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,

PT useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 11; Page 7367-7369; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG7788 represent human colon

XX cancer-associated nucleic acid molecules (N) and proteins (P), where

XX the proteins are collectively known as colon cancer antigens. The colon

XX cancer antigens have cytostatic activity and can be used in gene

XX therapy and vaccine production. N and P may be used in the prevention,

XX diagnosis and treatment of diseases associated with inappropriate P

XX expression. For example, N and P may be used to treat disorders

XX associated with decreased expression by rectifying mutations or deletions

XX in a patient's genome that affect the activity of P by expressing

XX inactive proteins or to supplement the patients own production of P.

XX Additionally, N may be used to produce the colon cancer-associated Ps,

XX by inserting the nucleic acids into a host cell and culturing the cell

XX to express the proteins. N and P can be used in the prevention, diagnosis

XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204

XX and AAB77789 represent sequences used in the exemplification of the

XX present invention.

XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were

XX missing at time of publication, meaning no sequences are present for

XX SEQ ID NO:1027 to 1052, 7921 and 7922.

XX SQ Sequence 544 AA;

Query Match 43.4%; Score 43; DB 22; Length 544;

Best Local Similarity 70.0%; Pred. No. 1.6e+02;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAKICHQIEY 13

Db 341 EALVCHQIRY 350

Search completed: April 23, 2003, 13:27:12

Job time : 28.5169 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:21:27 ; Search time 24.809 Seconds
(without alignments)
132.885 Million cell updates/sec

Title: US-09-836-073-19

Perfect score: 84

Sequence: 1 QERAIIRQVEYFGDF 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	92.9	390	5	Q8T8V5 drosophila
2	57	67.9	381	11	Q9CYB9 mus musculus
3	53	63.1	411	10	Q9FL36 Q9FI36 arabidopsis
4	53	63.1	422	10	Q9A338 Q9A338 arabidopsis
5	50	59.5	396	5	O01806 O01806 caenorhabdi
6	49	58.3	389	10	Q8S0T8 Q8S0T8 oryza sativ
7	48	57.1	433	10	Q932V7 Q932V7 arabidopsis
8	48	57.1	483	10	O65529 O65529 arabidopsis
9	47	56.0	206	13	Q8QH15 Q8QH15 gallus gall
10	47	56.0	545	10	O80567 O80567 arabidopsis
11	47	56.0	826	10	.Q940X9 Q940X9 arabidopsis
12	47	56.0	1379	5	Q9VAV5 Q9VAV5 drosophila
13	47	56.0	1403	5	Q9NH66 Q9NH66 drosophila
14	45.5	54.2	642	16	Q9KEU7 Q9KEU7 bacillus ha
15	45	53.6	914	12	Q85427 Q85427 rat cytochrome
16	43	51.2	343	2	Q9XBI7 Q9XBI7 bacillus ce

17	43	51.2	469	16	O66887
18	43	51.2	569	10	O49048
19	43	51.2	569	10	O80650
20	43	51.2	669	16	Q8XRQ1
21	43	51.2	846	16	Q9RT05
22	42	50.0	213	12	Q9YX11
23	42	50.0	296	11	Q9CTN3
24	42	50.0	343	10	Q94LD0
25	42	50.0	492	11	Q9D3J0
26	42	50.0	788	17	O58603
27	41.5	49.4	334	17	O59279
28	41	48.8	324	5	O76513
29	41	48.8	325	5	O17144
30	41	48.8	421	16	Q8YLF7
31	41	48.8	473	16	Q8XM38
32	41	48.8	483	5	Q19463
33	41	48.8	500	16	Q987T0
34	41	48.8	518	16	Q8RD16
35	41	48.8	520	16	Q9KNE3
36	40.5	48.2	928	10	Q9LU02
37	40.5	48.2	3078	5	Q26031
38	40	47.6	223	11	Q9R142
39	40	47.6	290	17	Q9HMC4
40	40	47.6	322	2	Q93P06
41	40	47.6	322	2	Q93P05
42	40	47.6	322	2	Q93P04
43	40	47.6	352	2	O07295
44	40	47.6	355	10	Q8RYB4
45	40	47.6	373	16	Q8XEV5

ALIGNMENTS

RESULT 1

Q8T8V5 Q8T8V5 PRELIMINARY; PRT; 390 AA.
AC Q8T8V5;
DT 01-JUN-2002 (TREMREL. 21, Created)
DT 01-JUN-2002 (TREMREL. 21, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE AT22034P.
GN LA.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacle J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celnikier S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY075257; AAL68124.1;
SQ SEQUENCE 390 AA; 44869 MW; 797FDE26B903C909 CRC64;

Query Match 92.9%; Score 78; DB 5; Length 390;
Best Local Similarity 100.0%; Pred. No. 4.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QERAIIRQVEYFGD 15

Db 50 QERAIIRQVEYFGD 64

RESULT 2

Q9CYB9 Q9CYB9 PRELIMINARY; PRT; 381 AA.
ID Q9CYB9
AC Q9CYB9;

DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Sjogren syndrome antigen B.
 GN SSB.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK017822; BAB30957.1; -
 DR MGI; MGI:98423; SSB.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;

 Query Match 67.9%; Score 57; DB 11; Length 381;
 Best Local Similarity 56.7%; Pred. No. 0.16;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

 QY 2 ERAIIOVEYFGDF 16
 Db 14 EAKICHQIEYFGDF 28

 RESULT 3
 Q9FL36 PRELIMINARY; PRT; 411 AA.
 AC Q9FL36;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Similarity to RNA-binding protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=98344145; PubMed=9679202;
 RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
 RT features of the regions of 1,381,565 bp covered by twenty one
 RT physically assigned P1 and TAC clones."
 RL DNA Res. 5:131-145(1998).

DR EMBL; AB010698; BAB11080.1; -
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 SQ SEQUENCE 411 AA; 45655 MW; A2EF62EB5B89B099 CRC64;

 Query Match 63.1%; Score 53; DB 10; Length 411;
 Best Local Similarity 90.9%; Pred. No. 0.83;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 5 IIRQVEYFGD 15
 Db 96 IIRQVEYFGD 106

 RESULT 4
 Q94A38 PRELIMINARY; PRT; 422 AA.
 ID Q94A38;
 AC Q94A38;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE AT5G46250/WPL12.3.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Arabidopsis cDNA clones."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY050403; AAK91419.1; -
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR PROSITE; PS0102; RRM; 1.
 SQ SEQUENCE 422 AA; 46842 MW; 4EC4BBBF1E068F0E CRC64;

 Query Match 63.1%; Score 53; DB 10; Length 422;
 Best Local Similarity 90.9%; Pred. No. 0.85;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

 QY 5 IIRQVEYFGD 15
 Db 107 IIRQVEYFGD 117

 RESULT 5
 O01806 PRELIMINARY; PRT; 396 AA.
 ID O01806;
 AC O01806;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE C44E4.4 protein.
 DE C44E4.4.
 GN C44E4.4.
 OS Caenorhabditis elegans.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.

```
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RL investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Sammons L., Wohldmann P., Gillam B.;
RT "The sequence of C. elegans cosmid C4E4.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003140; AAB54169.1; -.
DR InterPro; IPR002344; Lupus_La.
DR Pfam; PF00076; rim; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
SQ SEQUENCE 396 AA; 43631 MW; A2D828A4FAA3C34 CRC64;

Query Match 59.5%; Score 50; DB 5; Length 396;
Best Local Similarity 57.1%; Pred. No. 2.6;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERAIRQVEYYFGD 15
   ::::|::|::|
Db 13 DQKIQLQVEYFEGN 26

RESULT 6
Q8S0T8 PRELIMINARY; PRT; 389 AA.
AC Q8S0T8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE RNA-binding protein-like.
GN OJ1414.E05.3.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
RL clone:OJ1414.E05.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003375; BAB90138.1; -.
SQ SEQUENCE 389 AA; 42100 MW; 10C1863EAB6FA7B7 CRC64;

Query Match 58.3%; Score 49; DB 10; Length 389;
Best Local Similarity 72.7%; Pred. No. 3.7;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 IIRQVEYYFGD 15
   |::|::|::|::|
Db 91 IVKQVEYYFSD 101

RESULT 7
Q932V7 PRELIMINARY; PRT; 433 AA.
ID Q932V7
AC Q932V7;
```

```
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative RNA-binding protein LAH1.
GN AT4G32720.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full Length cDNA of gene AT4G32720 (GI:7270219).";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY056237; AAL07086.1; -.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rim; 1.
DR PROSITE; PS50102; RRM; 1.
SQ SEQUENCE 433 AA; 48095 MW; ES8EBAF51C35A8F7 CRC64;

Query Match 57.1%; Score 48; DB 10; Length 433;
Best Local Similarity 53.8%; Pred. No. 5.2;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 RAIIRQVEYYFGD 15
   ::|::|::|::|
Db 12 KTVLRQVEYFSD 24

RESULT 8
O65529 PRELIMINARY; PRT; 483 AA.
AC O65529;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 54.1 kDa protein.
GN F4D11.80 OR AT4G32720.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Benes V., Rechmann S., Borkova D., Ansonge W., Hoheisel J.,
RA Meves H.W., Mayer K.F.X., Schueller C.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Benes V., Rechmann S., Borkova D., Ansonge W., Meves H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL022537; CAAL8589.1; -.
DR EMBL; AL161582; CAB79989.1; -.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rim; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
KW Hypothetical protein.
SQ SEQUENCE 483 AA; 54125 MW; 20840B34A0BC94E2 CRC64;
```

```

Query Match          57.1%; Score 48; DB 10; Length 483;
Best Local Similarity 53.8%; Pred. No. 6.9;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 RAIIRQVEYFSD 15
DB 12 KTVLRQVEYFSD 24

RESULT 9
Q80H15 ID Q80H15 PRELIMINARY; PRT; 206 AA.
AC Q80H15;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Untranslated region binding-protein.
GN UBP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus
OX NCBI_TaxID=9031;
[1]
RN
RP SEQUENCE FROM N.A.
RA L'Ecuver T.J., Fang H.-L.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF467897; AAL76269.1; -;
SQ SEQUENCE 206 AA; 2392 MW; 965962f7dffb90e9 CRC64;

Query Match          56.0%; Score 47; DB 13; Length 206;
Best Local Similarity 57.1%; Pred. No. 4.2;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 ERAIRQVEYFSD 15
DB 14 ESKICQIEYFSGN 27

RESULT 10
O80567 ID O80567 PRELIMINARY; PRT; 545 AA.
AC O80567; Q9C5X1;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Expressed protein (VirF-interacting protein FIP1) (Hypothetical 60.6
DE kDa protein) (At2g43970/F6E13.10).
GN AT2G43970 OR F6E13.10/At2G43970.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN
RP SEQUENCE FROM N.A.
RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
RA Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
[4]

```

```

RP SEQUENCE FROM N.A.
RA Mayda E., Tzfira T., Citovsky V.;
RT "Arabidopsis thaliana VirF-interacting protein FIP1.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,
RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene F6E13.10/At2g43970 (GI:3212854).";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Koesema E., Meyers M.C., Shinn P.,
RA Banh J., Bowser L., Carninci P., Chung M.K., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yu G., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
[7]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004005; AAC23405.2; -;
DR EMBL; AF332565; AAK06847.1; -;
DR EMBL; AY056238; AAL07087.1; -;
DR EMBL; AF375410; AAK52994.1; -;
DR EMBL; AF367277; AAK56266.1; -;
DR InterPro: IPR002344; Lupus_La.
DR PRINTS; PR00302; LUPUSLA.
SQ Hypothetical protein.
KW SEQUENCE 545 AA; 60589 MW; E1A933261FE1ED80 CRC64;

Query Match          56.0%; Score 47; DB 10; Length 545;
Best Local Similarity 72.7%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIRQVEYFSD 15
DB 197 IVNQVEYFSD 207

RESULT 11
Q940X9 ID Q940X9 PRELIMINARY; PRT; 826 AA.
AC Q940X9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE AT5g21160/T10F18_190.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,

```


RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlín-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.C., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
RT "Arabidopsis cDNA clones."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY052365; AAK96556.1;
DR InterPro; IPR001950; TIF.SUI1.
DR PROSITE; PS01118; SUI1.1; UNKNOWN.1.
SQ SEQUENCE 826 AA; 91377 MW; 8D41922E5B609D9A CRC64;

Query Match 56.0%; Score 47; DB 10; Length 826;
Best Local Similarity 63.6%; Pred. No. 18;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
Db 282 VLKQVEYFSD 292

RESULT 12
Q9VAV5 PRELIMINARY; PRT; 1379 AA.

ID Q9VAV5
AC Q9VAV5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-WAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Larp protein.
GN LARP OR CG14066.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Abayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karp G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C., Kraft C., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacile J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AF003764; AAF56783.2;
DR FlyBase; FBgn0040108; larp.
SQ SEQUENCE 1379 AA; 147832 MW; EEA3E257A733641A CRC64;

Query Match 56.0%; Score 47; DB 5; Length 1379;
Best Local Similarity 52.4%; Pred. No. 30;
Matches 11; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

QY 2 ERAIRQVEYF-----GDF 16
Db 474 KEAIKQVEYFSDVNLGTGDF 494

RESULT 13
Q9NHN6 PRELIMINARY; PRT; 1403 AA.

ID Q9NHN6
AC Q9NHN6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE La related protein (Fragment).
GN LARP OR CG14066.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC Chauvet S., Maurel-Zaffran C., Miassod R.-M., Jullien N., Pradel J.,
RA Aragnol D.;
RT "Larp, a new candidate Hox target in Drosophila whose orthologue in
RT mouse is expressed at sites of epithelium/mesenchymal interactions."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF221108; AAF35862.1;
DR FlyBase; FBgn0040108; larp.
FT NON_TER 1
SQ SEQUENCE 1403 AA; 150920 MW; 047500A8CE901A38 CRC64;

Query Match 56.0%; Score 47; DB 5; Length 1403;
Best Local Similarity 52.4%; Pred. No. 31;
Matches 11; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

QY 2 ERAIRQVEYF-----GDF 16
Db 460 KEAIKQVEYFSDVNLGTGDF 480

RESULT 14
Q9KEU7 PRELIMINARY; PRT; 642 AA.

ID Q9KEU7
AC Q9KEU7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE ABC transporter (permease).
GN BH0752.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hiramata C., Nakamura Y., Ogasawara N., Kuhara S.,

Search completed: April 23, 2003, 13:33:01
Job time : 26.809 secs

RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RL halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL: AP001509; BAB04471.1;
DR InterPro: IPR003838; DUF214.
DR InterPro: IPR001991; Na/diCO_symp.
DR Pfam: PF02687; DUF214; 1.
DR PRINTS: PR00173; EDTRNSPORT.
KW Complete proteome.
SQ SEQUENCE 642 AA; 73354 MW; A44515A412PE61E9 CRC64;

Query Match 54.2%; Score 45.5; DB 16; Length 642;

Best Local Similarity 60.0%; Pred. No. 25;

Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 QERAIQVEVYFGD 15

||||: |||||

DB 471 QEKALVLH-EYVFGD 484

RESULT 15

Q85427

ID Q85427 PRELIMINARY; PRT; 914 AA.

AC Q85427;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE PR55.

GN R55.

OS Rat cytomegalovirus (strain Maastricht).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Muromegalovirus.

OX NCBI_TaxID=79700;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MAASTRICHT;

RX MEDLINE=96335691; PubMed=8757999;

RA Beuken E., Slobbe R., Bruggeman C.A., Vink C.;

RT "Cloning and sequence analysis of the genes encoding DNA polymerase,

glycoprotein B, gp18.5 and major DNA-binding protein of rat

cytomegalovirus.";

RL J. Gen. Virol. 77:1559-1562(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=MAASTRICHT;

RX MEDLINE=20366325; PubMed=10906222;

RA Vink C., Beuken E., Bruggeman C.A.;

RT "Complete DNA sequence of the rat cytomegalovirus genome.";

RL J. Virol. 74:7656-7663(2000).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=MAASTRICHT;

RX MEDLINE=20473137; PubMed=11018281;

RA Gruijthuisen Y.K., Beuken E., Bruggeman C.A., Vink C.;

RT "Rat cytomegalovirus R89 is a highly conserved gene which expresses a

spliced transcript.";

RL Virus Res. 59:119-130(2000).

DR EMBL: AF232689; AAC56432.1;

DR InterPro: IPR000234; Glycoprot_B.

DR InterPro: IPR000847; HTH_LysR.

DR Pfam: PF00606; Glycoprotein_B; 1.

DR ProDom: PD000693; Glycoprot_B; 1.

DR PROSITE: PS00044; HTH_LYSR_FAMILY; UNKNOWN_1.

SQ SEQUENCE 914 AA; 102697 MW; 74698190F2A2203A CRC64;

Query Match

Best Local Similarity 53.6%; Score 45; DB 12; Length 914;

Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 QERAIQVEVYF 13

||||: |||||

DB 792 QQAVMRPVEYFF 804

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:42 ; Search time 4.4049 Seconds
(without alignments)
150.669 Million cell updates/sec

Title: US-09-836-073-19
Perfect score: 84
Sequence: 1 QERAIKQVEYFGDF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	92.9	390	1 LA_DROME	P40796 drosophila
2	57	67.9	404	1 LA_BOVIN	P10881 bos taurus
3	57	67.9	408	1 LA_HUMAN	P05455 homo sapien
4	57	67.9	415	1 LA_MOUSE	P32067 mus musculus
5	57	67.9	415	1 LA_RAT	P38656 rattus norv
6	56	66.7	383	1 LA_AEDAL	Q28457 aedes albop
7	48	57.1	427	1 LAB_XENLA	P28049 xenopus lae
8	46	54.8	298	1 LAH1_SCHPO	P87058 schizosacch
9	44	52.4	315	1 HO2_RAT	P23711 rattus norv
10	44	52.4	315	1 HO2_RAT	P23711 rattus norv
11	43	51.2	466	1 SRO9_YEAST	P25567 saccharomyc
12	42	50.0	275	1 LAH1_YEAST	P33399 saccharomyc
13	42	50.0	398	1 YJGN_ECO57	P58219 escherichia
14	42	50.0	398	1 YJGN_ECOLI	P39338 escherichia
15	40.5	48.2	506	1 TCWO_PETCR	Q43033 petroselinu
16	40	47.6	304	1 Y687_HAEIN	P71356 haemophilus
17	40	47.6	373	1 TVRA_ECOLI	P07023 escherichia
18	40	47.6	634	1 IDUA_MOUSE	P48441 mus musculus
19	39	46.4	206	1 VADI_TREPA	O83443 treponema p
20	39	46.4	253	1 SOJ_TREPA	O83296 treponema p
21	39	46.4	267	1 PSTB_XYLFA	Q9pbk0 xylella fas
22	39	46.4	312	1 HO2_RABIT	P43242 oryctolagus
23	39	46.4	315	1 HO2_MOUSE	O70252 mus musculus
24	39	46.4	316	1 HO2_HUMAN	P30519 homo sapien
25	39	46.4	368	1 PFQO_LACDE	Q986S1 lactobacill
26	39	46.4	368	1 PFQO_LACDL	P46545 lactobacill
27	39	46.4	532	1 IPF7_SHIFL	P18014 shigella fl
28	39	46.4	602	1 EX5A_BUCAL	P57530 buchnera ap
29	38	45.2	625	1 EGAL_LACSK	Q48846 lactobacill
30	37	44.0	216	1 POLG_FMDVC	P03309 foot-and-mo
31	37	44.0	230	1 POLG_FMDV5	P03307 foot-and-mo
32	37	44.0	234	1 POLG_FMDVI	P03310 foot-and-mo
33	37	44.0	363	1 YCDM_ECO57	Q8xaul escherichia

RESULT 1

ID	LA_DROME	STANDARD;	PRT;	390 AA.
AC	P40796; Q24375; Q9VIN2;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	La protein homolog (La ribonucleoprotein) (La autoantigen homolog).			
GN	LA OR CGI0922.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;			
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Canton-S; TISSUE=Ovary;			
RX	MEDLINE=94309632; PubMed=8035794;			
RA	Bai C., Li Z., Tollas P.P.;			
RT	"Developmental characterization of a Drosophila RNA-binding protein homologous to the human systemic lupus erythematosus-associated La/SS-B autoantigen."			
RT	La/SS-B autoantigen."			
RL	Mol. Cell. Biol. 14:5123-5129(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94309661; PubMed=8035818;			
RA	Yoo C.J., Wolin S.L.;			
RT	"La proteins from Drosophila melanogaster and Saccharomyces cerevisiae: a yeast homolog of the La autoantigen is dispensable for growth."			
RT	Mol. Cell. Biol. 14:5412-5424(1994).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkley;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mklos G.L.G.,			
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			
RA	Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	De Pablos B., Delcher A., Deng X., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,			
RA	Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hoskin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			

34	37	44.0	363	1 YCDM_ECOLI	P75898 escherichia
35	37	44.0	550	1 SYR_MYCLE	P45840 mycobacteri
36	37	44.0	611	1 VATA_PLAFA	Q03498 plasmodium
37	37	44.0	781	1 APE2_SULTO	Q974n6 sulfolobus
38	37	44.0	900	1 GUNH_CLOTM	P16218 clostridium
39	37	44.0	903	1 CHO2_SCHPO	O74787 schizosacch
40	37	44.0	1011	1 POLG_FMDVT	P15072 foot-and-mo
41	37	44.0	1277	1 NPCL_PIG	P56941 sus scrofa
42	37	44.0	1912	1 CHD4_HUMAN	Q14839 homo sapien
43	37	44.0	1944	1 CHD3_HUMAN	Q12873 homo sapien
44	37	44.0	2095	1 BP28_DROME	Q9ym75 drosophila
45	37	44.0	2332	1 POLG_FMDVA	P03308 f genome po

ALIGNMENTS

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiner K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao O., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTION TERMINATION BY RNA
CC POLYMERASE III. BINDS RNA AND DNA. BINDS TO PRECURSORS OF RNA
CC POLYMERASE III TRANSCRIPTS. MAY PLAY A SPECIALIZED ROLE DURING FLY
CC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT EMBRYONIC, LARVAL,
CC PUPAL, AND ADULT DEVELOPMENT. EXPRESSION THROUGHOUT THE EMBRYO IS
CC FOLLOWED BY A RESTRICTED PATTERN OF MESODERMAL EXPRESSION THAT IS
CC LATER CONFINED TO THE VISCERAL MESODERM, GONADS, GUT, AND SALIVARY
CC GLANDS.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U07652; AAA20518.1; -
CC EMBL; L32988; AAA21776.1; -
CC EMBL; AE003666; AAF53885.1; -
CC EMBL; FB90011638; La.
CC InterPro; IPR002344; Lupus_La.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS00102; RRM; 1.
CC RNA-binding; Nuclear protein; DNA-binding.
CC DOMAIN 149 234 RNA-BINDING (RRM).
CC FT CONFLICT 169 169 A -> T (IN REF. 1).
CC FT CONFLICT 182 183 KH -> NS (IN REF. 1).
CC FT CONFLICT 283 283 A -> R (IN REF. 1).
CC FT CONFLICT 329 329 K -> N (IN REF. 1).
CC SQ SEQUENCE 390 AA; 44884 MW; A8099288B90446A5 CRC64;
Query Match 92.9%; Score 78; DB 1; Length 390;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 QERAIIRQVEYYFGD 15
Db 50 QERAIIRQVEYYFGD 64
RESULT 2
LA_BOVIN
ID LA_BOVIN STANDARD; PRT; 404 AA.
AC P10881.
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
DE homolog).
GN SSB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=99202037; PubMed=2468131;
RA Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
sequences for RNA-binding.";
RL Nucleic Acids Res. 17:2233-2244(1989).
CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
CC AND 7-2 RNAs.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
CC C-TERMINAL PART OF THE PROTEIN.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X13698; CAA31986.1; -
CC PIR; S03849; S03849.
CC InterPro; IPR002344; Lupus_La.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS00102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; 1.
CC RNA-binding; Nuclear protein; Phosphorylation.
CC FT DOMAIN 111 187 RNA-BINDING (RRM).
CC SQ SEQUENCE 404 AA; 46534 MW; 4EE30B5C262AD6A1 CRC64;
Query Match 67.9%; Score 57; DB 1; Length 404;
Best Local Similarity 66.7%; Pred. No. 0.013;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
OY 2 ERAIRQVEYYFGDF 16
Db 14 EAKICHQIEYYFGDF 28
RESULT 3
LA_HUMAN
ID LA_HUMAN STANDARD; PRT; 408 AA.
AC P03455;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lupus La protein (Sjogren syndrome type B antigen (SS-B)) (La
DE ribonucleoprotein) (La autoantigen).
GN SSB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE-89202037; PubMed-2468131;
 RA Chan E.K.L., Sullivan K.F., Tan E.M.;
 RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
 RT sequences for RNA-binding.";
 RN Nucleic Acids Res. 17:2233-2244 (1989).
 [2]
 RX SEQUENCE FROM N.A.
 RP MEDLINE-89053970; PubMed-3192525;
 RA Chambers J.C., Kenan D., Martin B.J., Keene J.D.;
 RT "Genomic structure and amino acid sequence domains of the human La
 RT autoantigen.";
 RN J. Biol. Chem. 263:18043-18051 (1988).
 [3]
 RX SEQUENCE OF 54-408 FROM N.A.
 RP MEDLINE-88199081; PubMed-2422201;
 RA Sturgess A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
 RA Coppel R.S.;
 RT "Characteristics and epitope mapping of a cloned human autoantigen
 RT La.";
 RN J. Immunol. 140:3212-3218 (1988).
 [4]
 RX SEQUENCE OF 54-97 FROM N.A.
 RP MEDLINE-85166283; PubMed-3856888;
 RA Chambers J.C., Keene J.D.;
 RT "Isolation and analysis of cDNA clones expressing human lupus La
 RT antigen.";
 RN Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119 (1985).
 [5]
 RX FUNCTION.
 RP MEDLINE-89251617; PubMed-2470590;
 RA Gottlieb E., Steitz J.A.;
 RT "Function of the mammalian La protein: evidence for its action in
 RT transcription termination by RNA polymerase III.";
 RN EMBO J. 8:851-861 (1989).
 [6]
 RX PHOSPHORYLATION.
 RP MEDLINE-97207017; PubMed-9054510;
 RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Marale R.J.;
 RT "Phosphorylation of the human La antigen on serine 366 can regulate
 RT recycling of RNA polymerase III transcription complexes.";
 RN Cell 88:707-715 (1997).
 CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
 CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
 CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
 CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
 CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
 CC AND 7-2 RNAS.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
 CC C-TERMINAL PART OF THE PROTEIN.
 CC -1- PTM: THE N-TERMINUS IS BLOCKED.
 CC -1- DISEASE: SERA FROM PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS
 CC OFTEN CONTAIN ANTIBODIES THAT REACT WITH THE NORMAL CELLULAR
 CC LA PROTEIN AS IF THESE ANTIGEN WAS FOREIGN.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X13697; CAA31985.1; -;
 DR EMBL: J04205; AAA31885.1; -;
 DR PIR: A31888; A31888.
 DR PIR: A22956; A22956.
 DR PIR: A31273; A31273.
 DR PIR: S03848; S03848.
 DR PIR: S11013; S11013.
 DR Genew: HGNC:11316; SSB.
 DR MIM: 109090; -;

DR InterPro: IPR002344; Lupus_La.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; rrm; 1.
 DR PRINTS; PRO0302; LUPUSLA.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 KW Systemic lupus erythematosus; RNA-binding; Phosphorylation;
 KW Nuclear protein.
 KW DOMAIN 111 187 RNA-BINDING (RRM).
 FT MOD_RES 366 366 PHOSPHORYLATION (BY CK2).
 SQ SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;

 Query Match 67.9%; Score 57; DB 1; Length 408;
 Best Local Similarity 66.7%; Pred. No. 0.014;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Caps 0;

 QY 2 ERAIRIOVEYFGDF 16
 DB 14 EAKICHOIEYFGDF 28

 RESULT 4
 LA_MOUSE STANDARD; PRT; 415 AA.
 ID LA_MOUSE
 AC F32087;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
 DE homolog).
 GN SSB OR SS-B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93203630; PubMed-8454877;
 RA Topfer F., Gordon T., McCluskey J.;
 RT "Characterization of the mouse autoantigen La (SS-B). Identification
 RT of conserved RNA-binding motifs, a putative ATP binding site and
 RT reactivity of recombinant protein with poly(U) and human
 RT autoantibodies.";
 RN J. Immunol. 150:3091-3100 (1993).
 [2]
 RP SEQUENCE OF 1-11 FROM N.A.
 RA Groelz D., Bachmann M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
 CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
 CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
 CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
 CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
 CC AND 7-2 RNAS.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L00993; CAA39415.1; -;
 DR EMBL: Y07951; CAA69249.1; -;
 DR MGD; MGI:98423; Ssb.
 DR InterPro: IPR002344; Lupus_La.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; rrm; 1.

DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS00102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP.1; 1.
 KW RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 187
 SQ SEQUENCE 415 AA; 47756 MW; 2D75197692FDC933 CRC64;

Query Match 67.9%; Score 57; DB 1; Length 415;
 Best Local Similarity 66.7%; Pred. No. 0.014;
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ERAIRQVEYFGDF 16
 DB 14 EAKICHQIEYFGDF 28

RESULT 5
 LA_RAT
 ID LA_RAT STANDARD; PRT; 415 AA.
 AC P38656;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 DE homology.
 GN SSB OR SS-B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=101116;
 RN [1]
 RP SEQUENCE FROM N.A.; PubMed=7916708;
 RX Sensel I., Troester H., Bartsch H., Schwemmler M., Igloi G.L., Bachmann M.;
 RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La: detection of species-specific variations.";
 RL Gene 126:265-268(1993).
 CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S, AND 7-2 RNAs.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X67859; CAA48043.1; ..
 DR PIR; JCI1494; JCI1494.
 DR InterPro; IPR002344; Lupus_La.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS00102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP.1; 1.
 KW RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 187
 SQ SEQUENCE 415 AA; 47777 MW; 033FD9CC1E475F98 CRC64;

Query Match 67.9%; Score 57; DB 1; Length 415;
 Best Local Similarity 66.7%; Pred. No. 0.014;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIRQVEYFGDF 16
 DB 14 EAKICHQIEYFGDF 28

RESULT 6
 LA_AEDAL
 ID LA_AEDAL STANDARD; PRT; 383 AA.
 AC Q26457;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 DE Aedes albopictus (Forest day mosquito).
 OS Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes.
 OX NCBI_TaxID=7160;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=96135233; PubMed=8551578;
 RA Pardigon N., Strauss J.H.;
 RT "Mosquito homolog of the La autoantigen binds to Sindbis virus RNA.";
 RL J. Virol. 70:1173-1181(1996).
 CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTION TERMINATION BY RNA POLYMERASE III. BINDS RNA AND DNA. BINDS TO THE 3' END OF THE MINUS STRAND OF SINDBIS VIRUS RNA. THIS MAY BE SIGNIFICANT FOR SINDBIS VIRUS RNA REPLICATION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. PRIMARILY NUCLEAR,, BUT SIGNIFICANT AMOUNTS ARE PRESENT IN THE CYTOPLASM.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S80954; AAB35931.1; ..
 DR InterPro; IPR002344; Lupus_La.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS00102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP.1; FALSE_NEG.
 KW RNA-binding; Nuclear protein; DNA-binding.
 FT DOMAIN 141 228
 SQ SEQUENCE 383 AA; 44430 MW; 4E5CC8F21C40F452 CRC64;

Query Match 66.7%; Score 56; DB 1; Length 383;
 Best Local Similarity 71.4%; Pred. No. 0.019;
 Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ERAIRQVEYFGD 15
 DB 44 EASTIRQLEYFGD 57

RESULT 7
 LAB_XENLA
 ID LAB_XENLA STANDARD; PRT; 427 AA.
 AC P28049;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen homolog B).
 DE homology.
 GN LAB1.

```
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte.
RX MEDLINE=93287095; PubMed=8510143;
RA Scherly D., Stutz F., Lin-Marq N., Clarkson S.G.;
RT "La proteins from Xenopus laevis. cDNA cloning and developmental
expression.";
RL J. Mol. Biol. 231:196-204(1993).
CC -|- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -|- DEVELOPMENTAL STAGE: BARELY DETECTABLE IN STAGE I/II OOCYTES,
CC ACCUMULATE IN STAGE III/IV OOCYTES, THEN EXHIBIT A ROUGHLY
CC CONSTANT STEADY STATE LEVEL IN MATURE OOCYTES, EGGS, AND EARLY
CC EMBRYOS.
CC -|- PTM: PHOSPHORYLATED (PROBABLE).
CC -|- MISCELLANEOUS: THERE ARE TWO FORMS OF LA, LAA AND LAB, IN XENOPUS.
CC -|- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
CC -|- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X6818; CAA48716.1; -
DR PIR; S28544; S28544.
DR PIR; S33817; S33817.
DR InterPro; IPR002344; Lupus_La.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
KW RNA-binding; Nuclear protein; Phosphorylation.
FT DOMAIN 110 202 RNA-BINDING (RRM).
FT DOMAIN 315 331 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 427 AA; 48995 MW; 45F3146F8934A355 CRC64;

Query Match 57.1%; Score 48; DB 1; Length 427;
Best Local Similarity 72.7%; Pred. No. 0.57;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
DB 16 ICEQIEYFGD 26

RESULT 8
LAA_XENLA STANDARD; PRT; 428 AA.
AC F28048;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lupus La protein homolog A (La ribonucleoprotein A) (La autoantigen
GN homolog A).
GN LAA1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
```

```
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=93287095; PubMed=8510143;
RA Scherly D., Stutz F., Lin-Marq N., Clarkson S.G.;
RT "La proteins from Xenopus laevis. cDNA cloning and developmental
expression.";
RL J. Mol. Biol. 231:196-204(1993).
CC -|- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -|- DEVELOPMENTAL STAGE: BARELY DETECTABLE IN STAGE I/II OOCYTES,
CC ACCUMULATE IN STAGE III/IV OOCYTES, THEN EXHIBIT A ROUGHLY
CC CONSTANT STEADY STATE LEVEL IN MATURE OOCYTES, EGGS, AND EARLY
CC EMBRYOS.
CC -|- PTM: PHOSPHORYLATED (PROBABLE).
CC -|- MISCELLANEOUS: THERE ARE TWO FORMS OF LA, LAA AND LAB, IN XENOPUS.
CC -|- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
CC -|- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X68817; CAA48715.1; -
DR PIR; S28545; S28545.
DR PIR; S33818; S33818.
DR InterPro; IPR002344; Lupus_La.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KW RNA-binding; Nuclear protein; Phosphorylation.
FT DOMAIN 111 203 RNA-BINDING (RRM).
FT DOMAIN 196 212 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 428 AA; 48864 MW; AEB3A38B7D2E3EC3 CRC64;

Query Match 57.1%; Score 48; DB 1; Length 428;
Best Local Similarity 72.7%; Pred. No. 0.57;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIRQVEYFGD 15
DB 17 ICEQIEYFGD 27

RESULT 9
LAHL_SCHPO STANDARD; PRT; 298 AA.
ID LAHL_SCHPO Q10458; Q13362;
AC P87058; Q10458; Q13362;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
GN SLA1 OR SPAC57A10.10C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98067398; PubMed=9404894;
```

64 EAENVLQVEYFSD 77

RESULT 10

HO2_RAT	STANDARD;	PRT;	315 AA.
ID	HO2_RAT		
AC	P23711;		
DT	01-NOV-1991 (Rel. 20, Created)		
DT	01-NOV-1991 (Rel. 20, Last sequence update)		
DT	30-MAY-2000 (Rel. 39, Last annotation update)		
DE	Heme oxygenase 2 (EC 1.14.99.3) (HO-2).		
GN	HMOX2.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
ON	NCBI_TaxID=101116;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Testis;		
RC	MEDLINE=90237051; PubMed=2185251;		
RX	Rotenberg M.O., Maines M.D.;		
RA	"Isolation, characterization, and expression in Escherichia coli of a		
RT	cDNA encoding rat heme oxygenase-2.";		
RL	J. Biol. Chem. 265:7501-7506(1990).		
RL	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Sprague-Dawley; TISSUE=Liver;		
RC	MEDLINE=94156193; PubMed=8112599;		
RX	McCoubrey W.K. Jr., Maines M.D.;		
RA	"The structure, organization and differential expression of the gene		
RT	encoding rat heme oxygenase-2.";		
RL	Gene 139:155-161(1994).		
RL	[3]		
RP	SEQUENCE OF 142-232 FROM N.A., AND PARTIAL SEQUENCE.		
RC	TISSUE=Testis, and Liver;		
RX	MEDLINE=88139412; PubMed=3343248;		
RA	Cruse I., Maines M.D.;		
RT	"Evidence suggesting that the two forms of heme oxygenase are		
RT	products of different genes.";		
RL	J. Biol. Chem. 263:3348-3353(1988).		
CC	-1- FUNCTION: HEME OXYGENASE CLEAVES THE HEME RING AT THE ALPHA		
CC	METHENE BRIDGE TO FORM BILIVERDIN. BILIVERDIN IS SUBSEQUENTLY		
CC	CONVERTED TO BILIRUBIN BY BILIVERDIN REDUCTASE. UNDER		
CC	PHYSIOLOGICAL CONDITIONS, THE ACTIVITY OF HEME OXYGENASE IS		
CC	HIGHEST IN THE SPLEEN, WHERE SENESCENT ERYTHROCYTES ARE		
CC	SEQUESTERED AND DESTROYED.		
CC	-1- FUNCTION: HEME OXYGENASE 2 COULD BE IMPLICATED IN THE PRODUCTION		
CC	OF CARBON MONOXIDE IN BRAIN WHERE IT COULD ACT AS A		
CC	NEUROTRANSMITTER.		
CC	-1- CATALYTIC ACTIVITY: Heme + 3 AH(2) + O(2) = biliverdin + Fe(2+) +		
CC	CO + 3 A + 3 H(2O).		
CC	-1- SUBCELLULAR LOCATION: Mitochondrial.		
CC	-1- TISSUE SPECIFICITY: WIDELY DISTRIBUTED IN BODY WITH A HIGH		
CC	CONCENTRATION IN THE BRAIN.		
CC	-1- INDUCTION: HEME OXYGENASE 2 ACTIVITY IS NON-INDUCIBLE.		
CC	-1- SIMILARITY: BELONGS TO THE HEME OXYGENASE FAMILY.		
CC	-1- SIMILARITY: CONTAINS 2 HEME REGULATORY MOTIFS (HRM).		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	at the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; J05405; AAA41340.1;		
DR	EMBL; U05013; AAA19130.1;		
DR	EMBL; M18918; AAA41347.1;		
DR	PIR; A29922; A29922.		
DR	PIR; A35199; A35199.		
DR	HSSP; P06762; IDVG.		
DR	InterPro; IPR002051; Heme_oxygenase.		

```

Query Match      54.8%; Score 46; DB 1; Length 298;
Best Local Similarity 50.0%; Pred. No. 0.86;
Matches 7; Conservative 4; Mismatches 3; Indels
QY      2 ERAIRQVEYYFCD 15
      |  ::::|::| |

```



```

Query Match          50.0%; Score 42; DB 1; Length 398;
Best Local Similarity 57.1%; Pred. NO. 6.1;
Matches      8; Conservative      2; Mismatches      4; Indels      0; Gaps      0;

QY      1 QERAIIRQVEYYFG 14
DB      276 QKMIITAQLIYYFG 289
      | | | | |
      | | | | |

RESULT 14
YJGN_ECOLI
IID YJGN_ECOLI STANDARD; PRT; 398 AA.
AC P39338; P76811; P39339;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yjgN.
GN YJGN OR B4257.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes.";
RL Nucleic Acids Res. 23:2105-2119(1995).
RN [2]
RP REVISIONS.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP CONCEPTUAL TRANSLATION.
RA Rudd K.E.;
RL Unpublished observations (JAN-2000).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT HAD TO
CC BE INTRODUCED IN POSITION 12 TO PRODUCE THIS ORF.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; U14003; AAA97153.1; ALT_FRAME.
DR DR ENBL; U14003; AAA97154.1; ALT_FRAME.
DR DR ENBL; AE000496; AAC77214.1; ALT_FRAME.
DR Ecogen; EG12533; yjgN.
CK KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 74 94 POTENTIAL.
FT TRANSMEM 96 116 POTENTIAL.
FT TRANSMEM 143 163 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 229 249 POTENTIAL.
FT TRANSMEM 279 299 POTENTIAL.
FT TRANSMEM 334 354 POTENTIAL.
SQ SEQUENCE 398 AA; 44877 MW; 814BFAD98E816A63 CRC64;

Query Match          50.0%; Score 42; DB 1; Length 398;

```

Best Local Similarity 57.1%; Pred. No. 6.1;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QERAIHQVEYFG 14
| : | | : | | |
DB 276 QRMIAQLIYFG 289

RESULT 15

TCMO_PETCR STANDARD; PRT; 506 AA.
ID TCMO_PETCR
AC Q43033;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Trans-cinnamate 4-monooxygenase (EC 1.14.13.11) (Cinnamic acid
4-hydroxylase) (CA4H) (C4H) (P450C4H) (Cytochrome P450 73).
GN CYP73A10 OR CYP73.
OS Petroselinum crispum (Parsley) (Petroselinum hortense).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
OX NCBI_TaxID=4043;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95320184; PubMed=7597051;
RA Logemann E., Parniske M., Hahlbrock K.;
RT "Modes of expression and common structural features of the complete
phenylalanine ammonia-lyase gene family in parsley.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:5905-5909(1995).
CC -1- FUNCTION: CONTROLS CARBON FLOW TO PIGMENTS ESSENTIAL FOR
POLLINATION OR UV PROTECTION, TO NUMEROUS PHYTOALEXINS SYNTHESIZED
BY PLANTS WHEN CHALLENGED BY PATHOGENS, AND TO LIGNINS.
CC -1- CATALYTIC ACTIVITY: Trans-cinnamate + NADPH + O(2) -> 4-
hydroxycinnamate + NADP(+) + H(2)O.
CC -1- PATHWAY: Phenylpropanoid metabolism; second step.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: L38898; AAC41660.1; -
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Heme; NADP.
FT BINDING 448 448 HEME (BY SIMILARITY).
SQ SEQUENCE 506 AA; 58047 MW; 32F00EE959D69CCF CRC64;

Query Match 48.2%; Score 40.5; DB 1; Length 506;
Best Local Similarity 56.2%; Pred. No. 15;
Matches 9; Conservative 3; Mismatches 3; Indels 1; Gaps 1;
QY 2 ERAIRQVVEYFGDF 16
| : | | : | | |
DB 213 ERSRLAQSFYHFGDF 228

Search completed: April 23, 2003, 13:28:13
Job time : 5.40449 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:03 ; Search time 8.9876 Seconds
(without alignments)
171.119 Million cell updates/sec

Title: US-09-836-073-19

Perfect score: 84

Sequence: 1 QERAIIRQVEYFGDF 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000.

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	92.9	390	2 A53773	La/SS-B homolog D-
2	78	92.9	390	2 A53781	ribonucleoprotein
3	57	67.9	404	1 S03849	ribonucleoprotein
4	57	67.9	408	1 A31888	ribonucleoprotein
5	57	67.9	415	1 JCI494	ribonucleoprotein
6	50	59.5	396	2 T30953	hypothetical prote
7	48	57.1	427	1 S33817	ribonucleoprotein
8	48	57.1	428	1 S33818	ribonucleoprotein
9	48	57.1	483	2 T04453	hypothetical prote
10	47	56.0	529	2 T00677	hypothetical prote
11	46	54.8	298	2 T38937	rna binding protei
12	46	54.8	298	2 T43542	RNA-binding protei
13	45.5	54.2	642	2 H83743	ABC transporter (p
14	44	52.4	315	1 A35199	heme oxygenase (de
15	43	51.2	466	2 S19365	hypothetical prote
16	43	51.2	469	2 C70357	hypothetical prote
17	43	51.2	569	2 T52056	vacuolar protein s
18	43	51.2	569	2 T00445	vacuolar protein-s
19	43	51.2	845	2 C75333	general secretion
20	42	50.0	275	2 B48600	RNA-binding protei
21	42	50.0	393	2 D65238	hypothetical 44.4
22	42	50.0	398	2 B91283	hypothetical prote
23	42	50.0	407	2 D86124	hypothetical prote
24	42	50.0	788	2 A71076	hypothetical prote
25	41.5	49.4	334	2 A71035	probable proteinas
26	41	48.8	421	2 AE2473	hypothetical prote
27	41	48.8	483	2 T20895	hypothetical prote
28	41	48.8	520	2 B82505	hypothetical prote
29	40.5	48.2	506	2 T14907	trans-cinnamate 4-

30 40.5 48.2 3078 2 T28432 variant-specific s
31 40 47.6 290 2 C84410 hypothetical prote
32 40 47.6 304 2 H64156 hypothetical prote
33 40 47.6 373 1 KMECTD chorismate mutase
34 40 47.6 373 2 G91061 chorismate mutase-
35 40 47.6 373 2 C85906 chorismate mutase-
36 40 47.6 373 2 A10832 prephenate dehydro
37 40 47.6 384 2 H96829 probable RNA-bind
38 40 47.6 391 2 H89777 capsular polysacch
39 40 47.6 634 1 A55683 L-iduronidase (EC
40 40 47.6 662 2 S55387 hemocyanin precurs
41 40 47.6 736 1 C69307 conserved hypothet
42 39.5 47.0 335 2 A75179 probable proteinas
43 39 46.4 206 2 A71326 probable V-type AT
44 39 46.4 253 2 B71346 probable Spo0J reg
45 39 46.4 267 2 D82593 phosphate ABC tran

ALIGNMENTS

RESULT 1

A53773

La/SS-B homolog D-la - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000

C:Accession: A53773

R;Bai, C.; Li, Z.; Tollas, P.P.

Mol. Cell. Biol. 14, 5123-5129, 1994

A:Title: Developmental characterization of a Drosophila RNA-binding protein homolo

A:Reference number: A53773; MUID:94309632; PMID:8035794

A:Accession: A53773

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-390 <BAI>

A:Cross-references: GB:U07652; NID:9464019; PIDN:AAA20518.1; PID:9464020

C:Genetics:

A:Gene: FlyBase:La

A:Cross-references: FlyBase:FBgn0011638

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: leucine zipper; RNA binding

Query Match 92.9%; Score 78; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 8.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QERAIIRQVEYFGD 15

DB 50 QERAIIRQVEYFGD 64

RESULT 2

A53781

ribonucleoprotein La - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 20-Sep-1999

C:Accession: A53781

R;Yoo, C.J.; Wolin, S.L.

Mol. Cell. Biol. 14, 5412-5424, 1994

A:Title: La proteins from Drosophila melanogaster and Saccharomyces cerevisiae: a

A:Reference number: A53781; MUID:94309661; PMID:8035818

A:Accession: A53781

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-390 <YOO>

A:Cross-references: GB:L32988; NID:9488469; PID:9488470

C:Genetics:

A:Gene: FlyBase:La

A:Cross-references: FlyBase:FBgn0011638

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: RNA binding

Query Match 92.9%; Score 78; DB 2; Length 390;

Best Local Similarity 100.0%; Pred. No. 8.1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QERAIIRQVEYFGD 15
| | | | | | | | | | | | | | | | |
Db 50 QERAIIRQVEYFGD 64

RESULT 3
S03849
ribonucleoprotein La - bovine
N;Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C;Accession: S03849
R;Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A;Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences
A;Reference number: S03848; MUID:89202037; PMID:2468131
A;Accession: S03849
A;Molecule type: mRNA
A;Residues: 1-404 <CHA>
A;Cross-references: EMBL:X13698; NID:g755; PIDN:CAA31986.1; PID:g756
A;Note: Part of this sequence was confirmed by protein sequencing
C;Comment: This protein associates with a variety of small RNA molecules, most of which
ay act as a transcription termination factor.
C;Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
C;Keywords: blocked amino end; phosphoprotein; RNA binding
F;112-178/Domain: ribonucleoprotein repeat homology <RRM>
F;113-118/Region: RNA-binding RNP2 motif
F;151-158/Region: RNA-binding RNP1 motif
F;228-404/Domain: phosphorylated status predicted <PHY>

Query Match 67.9%; Score 57; DB 1; Length 404;
Best Local Similarity 66.7%; Pred. No. 0.041;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERATIRQVEYFGD 16
| | | | | | | | | | | | | | | | |
Db 14 EAKTCHQIEYFGD 28

RESULT 4
A31888
ribonucleoprotein La - human
N;Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome antigen
C;Species: Homo sapiens (man)
C;Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C;Accession: A31888; S03848; A22956; A61051; S11013; I55553; I70205; I70206; A31273
R;Chambers, J.C.; Kenan, D.; Martin, B.J.; Keene, J.D.
J. Biol. Chem. 263, 18043-18051, 1988
A;Title: Genomic structure and amino acid sequence domains of the human La autoantigen.
A;Reference number: A31888; MUID:89053970; PMID:3192525
A;Accession: A31888
A;Molecule type: mRNA
A;Residues: 1-408 <CHA>
A;Cross-references: GB:J04205; NID:g178686; PIDN:AAA51885.1; PID:g178687
R;Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A;Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences
A;Reference number: S03848; MUID:89202037; PMID:2468131
A;Accession: S03848
A;Molecule type: mRNA
A;Residues: 1-408 <CH2>
A;Cross-references: EMBL:X13697; NID:g36414; PIDN:CAA31985.1; PID:g36415
R;Chambers, J.C.; Keene, J.D.
Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985
A;Title: Isolation and analysis of cDNA clones expressing human lupus La antigen.
A;Reference number: A22956; MUID:85166283; PMID:3856888
A;Accession: A22956
A;Molecule type: mRNA
A;Residues: 45-97, 'LK' <CH3>
A;Cross-references: GB:J04205

A;Note: this sequence has been revised in reference A31888
R;Nyman, U.; Ringertz, N.R.; Pettersson, I.
Immunol. Lett. 22, 65-72, 1989
A;Title: Demonstration of an amino terminal La epitope recognized by human anti-La s
A;Reference number: A61051; MUID:89379261; PMID:2476379
A;Accession: A61051
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-19, 'E', 21-47 <NYM>
R;Sturgess, A.D.; Peterson, M.G.; McNeillage, L.J.; Whittingham, S.; Coppel, R.L.
J. Immunol. 140, 3212-3218, 1988
A;Title: Characteristics and epitope mapping of a cloned human autoantigen La.
A;Reference number: S11013; MUID:88199081; PMID:2452201
A;Accession: S11013
A;Molecule type: mRNA
A;Residues: 'E', 55-287, 'V', 289-408 <STU>
A;Cross-references: EMBL:M20328; NID:g337456; PIDN:AAA36577.1; PID:g337457
R;Kohsaka, H.; Yamamoto, K.; Fujii, H.; Miura, H.; Miyasaka, N.; Nishioaka, K.; Miyam
J. Clin. Invest. 85, 1566-1574, 1990
A;Title: Fine epitope mapping the human SS-B/La protein: Identification of a distinct
A;Reference number: I55553; MUID:90237237; PMID:1692037
A;Accession: I55553
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 81-107 <RES>
A;Cross-references: GB:M35261; NID:g338491; PIDN:AAA36652.1; PID:g338495
A;Accession: I70205
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 174-224 <RE2>
A;Cross-references: GB:M35263; NID:g338492; PIDN:AAA36653.1; PID:g338496
A;Accession: I70206
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 279-342 <RE3>
A;Cross-references: GB:M35262; NID:g338493; PIDN:AAA36654.1; PID:g338497
C;Comment: This protein associates with a variety of small RNA molecules, most of wh
ay act as a transcription termination factor.
C;Genetics:
A;Gene: GDB:SSB
A;Cross-references: GDB:125359; OMIM:109090
A;Map position: 2
A;Introns: 22/3; 57/2; 115/3; 185/2; 209/2; 223/3; 264/3; 380/2
C;Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
C;Keywords: phosphoprotein; RNA binding
F;112-178/Domain: ribonucleoprotein repeat homology <RRM>
F;113-118/Region: RNA-binding RNP2 motif
F;151-158/Region: RNA-binding RNP1 motif
F;228-408/Domain: phosphorylated status experimental <PHY>

Query Match 67.9%; Score 57; DB 1; Length 408;
Best Local Similarity 66.7%; Pred. No. 0.041;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERATIRQVEYFGD 16
| | | | | | | | | | | | | | | | |
Db 14 EAKTCHQIEYFGD 28

RESULT 5
JC1494
ribonucleoprotein La - rat
N;Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C;Species: Rattus norvegicus (Norway rat)
C;Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C;Accession: JC1494; S25145
R;Samsai, I.; Troester, H.; Bartsch, H.; Schwemmler, M.; Igloi, G.L.; Bachmann, M.
Gene 126, 265-268, 1993
A;Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection
A;Reference number: JC1494; MUID:93246255; PMID:7916708
A;Accession: JC1494
A;Molecule type: mRNA
A;Residues: 1-415 <SEM>

A:Cross-references: GB:X67859; NID:g55778; PIDN:CAA48043.1; PID:g55779
 A:Experimental source: liver
 C:Comment: This protein associates with a variety of small RNA molecules, most of which
 ay act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:113-178/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:227-415/Domain: phosphorylated #status predicted <PHY>

Query Match 67.9%; Score 57; DB 1; Length 415;
 Best Local Similarity 66.7%; Pred. No. 0.042;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIIQVEYYFGDF 16
 | | | | | | | | | |
 Db 14 EAKICHQIEYYFGDF 28

RESULT 6
 T30953
 hypothetical protein C44E4.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
 C:Accession: T30953
 R:Sammons, L.; Wohldmann, P.; Gillam, B.
 submitted to the EMBL Data Library, August 1999
 A:Description: The sequence of C. elegans cosmid C44E4.
 A:Reference number: Z20945
 A:Accession: T30953
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-396 <SAM>
 A:Cross-references: EMBL:AF003140; PIDN:AAB54169.1
 A:Experimental source: strain Bristol N2; clone C44E4
 C:Genetics:
 A:Map position: I
 A:Introns: 45/1; 114/3
 A:Note: C44E4.4
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

Query Match 59.5%; Score 50; DB 2; Length 396;
 Best Local Similarity 57.1%; Pred. No. 0.68;
 Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERAIIQVEYYFGD 15
 ::|||:|||||:
 Db 13 DOKIIQLEYYFGN 26

RESULT 7
 S33817
 ribonucleoprotein La.B - African clawed frog
 N:Alternate names: autoantigen SS-B/la; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33817; S28544
 R:Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression.
 A:Reference number: S33817; MUID:93287095; PMID:8510143
 A:Accession: S33817
 A:Molecule type: mRNA
 A:Residues: 1-427 <SCH>
 A:Cross-references: EMBL:X68818; NID:g64875; PIDN:CAA48716.1; PID:g64876
 C:Comment: This protein associates with a variety of small RNA molecules, most of which
 ay act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:111-177/Domain: ribonucleoprotein repeat homology <RRM>
 F:112-117/Region: RNA-binding RNP2 motif
 F:150-157/Region: RNA-binding RNP1 motif

F:227-427/Domain: phosphorylated #status predicted <PHY>

Query Match 57.1%; Score 48; DB 1; Length 427;
 Best Local Similarity 72.7%; Pred. No. 1.7;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIRQVEYYFGD 15
 | | | | | | | | | |
 Db 16 ICEQIEYYFGD 26

RESULT 8
 S33818
 ribonucleoprotein La.A - African clawed frog
 N:Alternate names: autoantigen SS-B/la; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33818; S28545
 R:Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression
 A:Reference number: S33817; MUID:93287095; PMID:8510143
 A:Accession: S33818
 A:Molecule type: mRNA
 A:Residues: 1-428 <SCH>
 A:Cross-references: EMBL:X68817; NID:g64873; PIDN:CAA48715.1; PID:g64874
 C:Comment: This protein associates with a variety of small RNA molecules, most of w
 ay act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:228-428/Domain: phosphorylated #status predicted <PHY>

Query Match 57.1%; Score 48; DB 1; Length 428;
 Best Local Similarity 72.7%; Pred. No. 1.7;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIRQVEYYFGD 15
 | | | | | | | | | |
 Db 17 ICEQIEYYFGD 27

RESULT 9
 T04453
 hypothetical protein F4D11.80 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
 C:Accession: T04453
 R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Anstorge, W.; Hohelsel, J.; Mewes
 submitted to the Protein Sequence Database, April 1998
 A:Reference number: Z15360
 A:Accession: T04453
 A:Molecule type: DNA
 A:Residues: 1-483 <BEV>
 A:Cross-references: EMBL:AL022537
 A:Experimental source: cultivar Columbia; BAC clone F4D11
 C:Genetics:
 A:Map position: 4
 A:Introns: 17/3; 44/1; 94/1; 141/3; 206/2; 247/2; 345/3; 395/1; 434/2
 A:Note: F4D11.80

Query Match 57.1%; Score 48; DB 2; Length 483;
 Best Local Similarity 53.8%; Pred. No. 1.9;
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 RAIIRQVEYYFGD 15
 : : | | | | : | |
 Db 12 KTVLRQVEYFSD 24

RESULT 10

T00677

hypothetical protein At2g43970 [imported] - Arabidopsis thaliana

N;Alternate names: hypothetical protein F6E13.10

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001

C;Accession: T00677; G84872

R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,

submitted to the EMBL Data Library, June 1998

A;Description: Arabidopsis thaliana chromosome II BAC F6E13 genomic sequence.

A;Reference number: Z14180

A;Accession: T00677

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-529 <ROU>

A;Cross-references: EMBL:AC004005; NID:g3212846; PID:g3212854

A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: G84872

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-529 <STO>

A;Cross-references: GB:AE002093; NID:g3212854; PIDN:AAC23405.1; GSPDB:GN00139

C;Genetics:

A;Gene: F6E13.10; At2g43970

A;Map position: 2

A;Introns: 200/3; 228/1; 259/3; 283/3; 310/2; 344/3; 367/3; 421/3

Query Match

Best Local Similarity 56.0%; Score 47; DB 2; Length 529;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 IIRQVEYYFGD 15

I: |||||

Db 197 IVNQVEYFSD 207

RESULT 11

T38937

rna binding protein - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C;Accession: T38937

R;Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, April 1997

A;Reference number: Z21818

A;Accession: T38937

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-298 <BAD>

A;Cross-references: EMBL:Z94864; PIDN:CA808173.1; GSPDB:GN00066; SPDB:SPAC57A10.10c

A;Experimental source: strain 972h-; cosmid c57A10

C;Genetics:

A;Gene: SPDB:SPAC57A10.10c

A;Map position: 1

A;Introns: 72/1

Query Match

Best Local Similarity 54.8%; Score 46; DB 2; Length 298;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ERAIRQVEYYFGD 15

I: |||||

Db 64 EAEVLKQVEYFSD 77

RESULT 12

T43542

RNA-binding protein Lal homolog - fission yeast (Schizosaccharomyces pombe)

N;Alternate names: La autoantigen; ribonucleoprotein La homolog

C;Species: Schizosaccharomyces pombe

C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

C;Accession: T43542; T43325

R;Van Horn, D.J.; Yoo, C.J.; Xue, D.; Shi, H.; Wolin, S.L.

RNA 3, 1434-1443, 1997

A;Title: The La protein in Schizosaccharomyces pombe: a conserved yet dispensable pr

A;Reference number: 222560; MUID:98067398; PMID:9404894

A;Accession: T43542

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-298 <VAN>

A;Cross-references: EMBL:AF022949; PIDN:AAB82145.1

R;Utsumi, R.

submitted to the EMBL Data Library, February 1998

A;Description: Screening of S. pombe cDNA library using E. coli defective in signal

A;Reference number: 222428

A;Accession: T43325

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-298 <UTS>

A;Cross-references: EMBL:AB011371; PIDN:BA24981.1

C;Genetics:

A;Gene: slal

C;Function:

A;Description: the binding of the La protein to tRNA precursors is required for the

C;Keywords: phosphoprotein; RNA binding

Query Match

Best Local Similarity 54.8%; Score 46; DB 2; Length 298;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ERAIRQVEYYFGD 15

I: |||||

Db 64 EAEVLKQVEYFSD 77

RESULT 13

H83743

ABC transporter (permease) BH0752 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C;Accession: H83743

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.;

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: H83743

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-642 <STO>

A;Cross-references: GB:AF001509; GB:BA000004; NID:gl0173176; PIDN:BA804471.1; GSPDB

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH0752

Query Match

Best Local Similarity 54.2%; Score 45.5; DB 2; Length 642;

Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Qy 1 QERAIQVEYYFGD 15

I: |||||

Db 471 QEKALVLH-EYYFGD 484

RESULT 14

A35199

heme oxygenase (decyclizing) (EC 1.14.99.3) 2 [similarity] - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C;Accession: A35199; A29922

R;Rotenberg, M.O.; Maines, M.D.

J. Biol. Chem. 265, 7501-7506, 1990

A;Title: Isolation, characterization, and expression in *Escherichia coli* of a cDNA encoding
A;Reference number: A35199; MUID:90237051; PMID:2185251
A;Accession: A35199
A;Molecule type: mRNA
A;Residues: 1-315 <ROT>
A;Cross-references: GB:J05405; NID:g204626; PIDN:AAA41340.1; PID:g204627
R;Cruse, I.; Maine, M.D.
J. Biol. Chem. 263, 3348-3353, 1988

A;Title: Evidence suggesting that the two forms of heme oxygenase are products of different
A;Reference number: A29922; MUID:88139412; PMID:3343248
A;Accession: A29922
A;Molecule type: mRNA
A;Residues: 147-229, 'TER' <CRU>
A;Cross-references: GB:M18918; NID:g204649; PIDN:AAA41347.1; PID:g554443
C;Superfamily: heme oxygenase (decyclizing)
C;Keywords: oxidoreductase

Query Match 52.4%; Score 44; DB 1; Length 315;
Best Local Similarity 40.0%; Pred. No. 6;
Matches 6; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Oy 1 QERAIROVEYFGD 15
Db 104 RREALIKMEYFFGE 118

RESULT 15
S19365
hypothetical protein YCL037c - yeast (*Saccharomyces cerevisiae*)
C;Species: *Saccharomyces cerevisiae*
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Jun-2000
C;Accession: S19365
R;Delgado, M.; Esteban, M.; Navas, L.
submitted to the Protein Sequence Database, March 1992
A;Reference number: S19365
A;Accession: S19365
A;Molecule type: DNA
A;Residues: 1-466
A;Cross-references: EMBL:X59720; NID:gl907116; PIDN:CAA42379.1; PID:g5326; MIPS:YCL037c
C;Genetics:
A;Gene: SGD:SR09
A;Cross-references: SGD:S0000542; MIPS:YCL037c
A;Map position: 3L

Query Match 51.2%; Score 43; DB 2; Length 466;
Best Local Similarity 53.6%; Pred. No. 14;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 5 IIRQVEYFGD 15
Db 297 IARQIEYFSE 307

Search completed: April 23, 2003, 13:34:44
Job time : 10.9888 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:27:28 ; Search time 9.16854 Seconds
(without alignments)
139.834 Million cell updates/sec

Title: US-09-836-073-19

Perfect score: 84
Sequence: 1 QERAIROVEYFGDF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PublishedApplications_AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	84	100.0	16	9 US-09-836-073-19	Sequence 19, Appl
2	62	73.8	18	9 US-09-836-073-4	Sequence 4, Appl
3	58	69.0	18	9 US-09-836-073-2	Sequence 2, Appl
4	57	67.9	17	9 US-09-836-073-13	Sequence 13, Appl
5	57	67.9	18	9 US-09-836-073-1	Sequence 1, Appl
6	57	67.9	18	9 US-09-836-073-5	Sequence 5, Appl
7	57	67.9	18	9 US-09-836-073-14	Sequence 14, Appl
8	57	67.9	39	9 US-09-843-676-26	Sequence 26, Appl
9	57	67.9	39	9 US-09-766-253-26	Sequence 26, Appl
10	57	67.9	39	9 US-09-438-486-26	Sequence 26, Appl
11	57	67.9	39	9 US-10-053-758-26	Sequence 26, Appl
12	57	67.9	39	9 US-10-054-295-26	Sequence 26, Appl
13	57	67.9	39	9 US-10-054-611-26	Sequence 26, Appl
14	57	67.9	460	9 US-10-102-806-695	Sequence 695, App
15	54	64.3	18	9 US-09-836-073-15	Sequence 15, Appl
16	51	60.7	18	9 US-09-836-073-9	Sequence 9, Appl
17	50	59.5	18	9 US-09-836-073-17	Sequence 17, Appl
18	49	58.3	18	9 US-09-836-073-11	Sequence 11, Appl
19	49	58.3	18	9 US-09-836-073-12	Sequence 12, Appl

20	48	57.1	18	9 US-09-836-073-10	Sequence 10, Appl
21	48	57.1	38	9 US-09-843-676-25	Sequence 25, Appl
22	48	57.1	38	9 US-09-766-253-25	Sequence 25, Appl
23	48	57.1	38	9 US-09-438-486-25	Sequence 25, Appl
24	48	57.1	38	9 US-10-053-758-25	Sequence 25, Appl
25	48	57.1	38	9 US-10-054-295-25	Sequence 25, Appl
26	48	57.1	38	9 US-10-054-611-25	Sequence 25, Appl
27	46.5	55.4	19	9 US-09-836-073-16	Sequence 16, Appl
28	45	53.6	18	9 US-09-836-073-3	Sequence 3, Appl
29	44.5	53.0	18	9 US-09-836-073-18	Sequence 18, Appl
30	43	51.2	37	9 US-09-843-676-24	Sequence 24, Appl
31	43	51.2	37	9 US-09-766-253-24	Sequence 24, Appl
32	43	51.2	37	9 US-09-438-486-24	Sequence 24, Appl
33	43	51.2	37	9 US-10-053-758-24	Sequence 24, Appl
34	43	51.2	37	9 US-10-054-295-24	Sequence 24, Appl
35	43	51.2	37	9 US-10-054-611-24	Sequence 24, Appl
36	42	50.0	18	9 US-09-836-073-7	Sequence 7, Appl
37	42	50.0	38	9 US-09-843-676-27	Sequence 27, Appl
38	42	50.0	38	9 US-09-766-253-27	Sequence 27, Appl
39	42	50.0	38	9 US-09-438-486-27	Sequence 27, Appl
40	42	50.0	38	9 US-10-053-758-27	Sequence 27, Appl
41	42	50.0	38	9 US-10-054-295-27	Sequence 27, Appl
42	42	50.0	38	9 US-10-054-611-27	Sequence 27, Appl
43	41	48.8	18	9 US-09-836-073-8	Sequence 8, Appl
44	40.5	48.2	2710	9 US-10-153-273-12	Sequence 12, Appl
45	40	47.6	223	12 US-10-005-168-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-836-073-19
; Sequence 19, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baldya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila
US-09-836-073-19

Query Match 100.0%; Score 84; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 3,7e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QERAIROVEYFGDF 16
Db 1 QERAIROVEYFGDF 16

RESULT 2
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baldya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match 73.8%; Score 62; DB 9; Length 18;
Best Local Similarity 68.8%; Pred. No. 0.00024;
Matches 11; Conservative 1; Mismatches 4; Indels 0;

QY 1 QERAIHQVEYFEGDF 16
||| :|||||
Db 3 QEAKICHQIEYFEGDF 18

RESULT 3

US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match 69.0%; Score 58; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.0012;
Matches 10; Conservative 2; Mismatches 3; Indels 0;

QY 2 ERAIRQVEYFEGDF 16
||| :|||||
Db 4 EAQICQIEYFEGDF 18

RESULT 4

US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match 67.9%; Score 57; DB 9; Length 17;
Best Local Similarity 66.7%; Pred. No. 0.0016;
Matches 10; Conservative 1; Mismatches 4; Indels 0;

QY 2 ERAIRQVEYFEGDF 16
||| :|||||
Db 3 EAKICHQIEYFEGDF 17

RESULT 5

US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match 67.9%; Score 57; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.0017;
Matches 10; Conservative 1; Mismatches 4; Indels 0;

QY 2 ERAIRQVEYFEGDF 16
||| :|||||
Db 4 EAKICHQIEYFEGDF 18

RESULT 6

US-09-836-073-5
; Sequence 5, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-5

Query Match 67.9%; Score 57; DB 9; Length 18;
Best Local Similarity 62.5%; Pred. No. 0.0017;
Matches 10; Conservative 2; Mismatches 4; Indels 0;

QY 1 QERAIHQVEYFEGDF 16
||| :|||||
Db 3 QEQKCHQIEYFEGDF 18

RESULT 7

US-09-836-073-14

; Sequence 14, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 22002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; PRIOR FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Bovine
 ; US-09-836-073-14

Query Match 67.9%; Score 57; DB 9; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.0017;
 Matches 10; Conservative 1; Mismatches 4; Indels 0;

QY 2 ERAIRQVEYFQDF 16
 | | | | | | | | | | | | | | | | | |
 Db 4 EAKICHOIEYFQDF 18

RESULT 8
 US-09-843-676-26
 ; Sequence 26, Application US/09843676
 ; Patent No. US20020164786A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cech, Thomas R.
 ; Lingner, Joachim
 ; Nakamura, Toru
 ; Chapman, Karen B.
 ; Morin, Gregg B.
 ; Harley, Calvin
 ; Andrews, William H.
 ; TITLE OF INVENTION: No. US20020164786A1el Telomerase
 ; NUMBER OF SEQUENCES: 225
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, 8th Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 94111

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/843,676
 FILING DATE: 26-Apr-2001
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/854,050
 FILING DATE: 09-MAY-1997
 APPLICATION NUMBER: US 08/846,017
 FILING DATE: 25-APR-1997
 APPLICATION NUMBER: US 08/844,419
 FILING DATE: 18-APR-1997
 APPLICATION NUMBER: US 08/724,643
 FILING DATE: 01-OCT-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Apple, Randolph T.
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 015389-002930US
 TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 39 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: No. US20020164786A1 Relevant
 ; TOPOLOGY: No. US20020164786A1 Relevant
 ; MOLECULE TYPE: peptide
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
 ; US-09-843-676-26

Query Match 67.9%; Score 57; DB 9; Length 39;
 Best Local Similarity 90.9%; Pred. No. 0.0039;
 Matches 10; Conservative 1; Mismatches 0; Indels 0;

QY 5 IIRQVEYFQDF 15
 | | | | | | | | | | | | | | | | | |
 Db 1 ILRQVEYFQDF 11

RESULT 9
 US-09-766-253-26
 ; Sequence 26, Application US/09766253
 ; Publication No. US20020187471A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cech, Thomas R.
 ; Lingner, Joachim
 ; Nakamura, Toru
 ; Chapman, Karen B.
 ; Morin, Gregg B.
 ; Harley, Calvin
 ; Andrews, William H.
 ; TITLE OF INVENTION: No. US20020187471A1el Telomerase
 ; NUMBER OF SEQUENCES: 171
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, 8th Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 94111

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/766,253
 FILING DATE: 19-Jan-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/846,017
 FILING DATE: 1997-04-25
 APPLICATION NUMBER: US 08/724,643
 FILING DATE: 01-OCT-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Apple, Randolph T.
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 015389-002920US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 39 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: not relevant
 ; MOLECULE TYPE: peptide
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
 ; US-09-766-253-26

Query Match 67.9% Score 57; DB 9; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.0039;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 ILRQVEYIFGD 15
I:|||||
Db 1 ILRQVEYIFGD 11

RESULT 10
US-09-438-486-26
; Sequence 26, Application US/09438486
; Publication No. US20030009019A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; LINGNER, Joachim
; NAKAMURA, Toru
; APPLICANT: NAKAMURA, Toru
; APPLICANT: CHAPMAN, Karen B.
; APPLICANT: MORIN, Gregg B.
; APPLICANT: HARLEY, Calvin
; APPLICANT: ANDREWS, William H.
; TITLE OF INVENTION: No. US20030009019A1 Telomerase
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/438,486
; FILING DATE: 12-NOV-1999
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002931US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-09-438-486-26

Query Match 67.9% Score 57; DB 9; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.0039;

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 5 ILRQVEYIFGD 15
I:|||||
Db 1 ILRQVEYIFGD 11

RESULT 11
US-10-053-758-26
; Sequence 26, Application US/10053758
; Publication No. US20030032075A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; LINGNER, Joachim
; NAKAMURA, Toru
; APPLICANT: CHAPMAN, Karen B.
; APPLICANT: MORIN, Gregg B.
; APPLICANT: HARLEY, Calvin
; APPLICANT: ANDREWS, William H.
; TITLE OF INVENTION: No. US20030032075A1 Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/053,758
; FILING DATE: 18-Jan-2002
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030032075A1 Relevant
; TOPOLOGY: No. US20030032075A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-10-053-758-26

Query Match 67.9% Score 57; DB 9; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.0039;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 ILRQVEYIFGD 15
I:|||||
Db 1 ILRQVEYIFGD 11

RESULT 12

US-10-054-295-26

Sequence 26, Application US/10054295

Publication No. US20030044953A1

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin

Andrews, William H.

TITLE OF INVENTION: No. US20030044953A1 Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/054,295

FILING DATE: 18-Jan-2002

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/854,050

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0300

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 39 amino acids

TYPE: amino acid

STRANDEDNESS: No. US20030044953A1 Relevant

TOPOLOGY: No. US20030044953A1 Relevant

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-10-054-295-26

Query Match

Best Local Similarity 67.9%; Score 57; DB 9; Length 39;

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 ILRQVEYFGD 15

Db 1 ILRQVEYFGD 11

RESULT 13

US-10-054-611-26

Sequence 26, Application US/10054611

Publication No. US20030059787A1

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin

Andrews, William H.

TITLE OF INVENTION: No. US20030059787A1 Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/054,611

FILING DATE: 18-Jan-2002

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/854,050

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0300

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 39 amino acids

TYPE: amino acid

STRANDEDNESS: No. US20030059787A1 Relevant

TOPOLOGY: No. US20030059787A1 Relevant

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-10-054-611-26

Query Match

Best Local Similarity 67.9%; Score 57; DB 9; Length 39;

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 ILRQVEYFGD 15

Db 1 ILRQVEYFGD 11

RESULT 14

US-10-102-806-695

Sequence 695, Application US/10102806

Publication No. US20030054421A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA103P1C1

CURRENT APPLICATION NUMBER: US/10/102,806

CURRENT FILING DATE: 2002-03-22

PRIOR APPLICATION NUMBER: 09/925,298

PRIOR FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05881

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 695
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-102-806-695

Query Match 67.9%; Score 57; DB 9; Length 460;
Best Local Similarity 66.7%; Pred. No. 0.05;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 ERAIIRQVEYYFGDF 16
| | | | | | | | | |
Db 66 EAKICHQIEYYFGDF 80

RESULT 15

US-09-836-073-15
; Sequence 15, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836.073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Xenopus
US-09-836-073-15

Query Match 64.3%; Score 54; DB 9; Length 18;
Best Local Similarity 75.0%; Pred. No. 0.0057;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 IIRQVEYYFGDF 16
| | | | | | | | | |
Db 7 ICEQIEYYFGDF 18

Search completed: April 23, 2003, 13:38:23
Job time : 9.16854 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:24 ; Search time 9.97753 Seconds
(without alignments)
47.183 Million cell updates/sec

Title: US-09-836-073-19

Perfect score: 84
Sequence: 1 QERAIIRQVEYFGDF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pap.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	67.9	18	4	US-09-316-630-3
2	57	67.9	18	4	US-09-316-630-4
3	57	67.9	39	3	US-08-851-843A-26
4	57	67.9	39	4	US-08-974-549A-216
5	57	67.9	39	4	US-08-854-050-26
6	57	67.9	39	4	US-09-430-323-26
7	50	59.5	38	4	US-08-974-549A-214
8	48	57.1	38	3	US-08-851-843A-25
9	48	57.1	38	4	US-08-974-549A-215
10	48	57.1	38	4	US-08-854-050-25
11	48	57.1	38	4	US-09-430-323-25
12	43	51.2	37	3	US-08-851-843A-24
13	43	51.2	37	4	US-08-854-050-24
14	43	51.2	37	4	US-09-430-323-24
15	42	50.0	38	3	US-08-851-843A-27
16	42	50.0	38	4	US-08-974-549A-217
17	42	50.0	38	4	US-08-854-050-27
18	42	50.0	38	4	US-09-430-323-27
19	40.5	48.2	2710	2	US-08-568-459A-12
20	40.5	48.2	2710	2	US-08-487-826B-12
21	40.5	48.2	2710	4	US-09-210-288-12
22	40.5	48.2	3060	2	US-08-487-826B-14
23	40	47.6	1076	4	US-09-470-443-6
24	40	47.6	1145	4	US-09-470-443-2
25	40	47.6	1145	4	US-09-470-443-4
26	39	46.4	432	4	US-09-282-305-10
27	38	45.2	166	4	US-09-199-637A-245

28	37.5	44.6	501	4	US-09-158-767-15	Sequence 15, Appl
29	37.5	44.6	501	4	US-09-158-767-16	Sequence 16, Appl
30	37.5	44.6	501	4	US-09-158-767-17	Sequence 17, Appl
31	37.5	44.6	501	4	US-09-158-767-18	Sequence 18, Appl
32	37	44.0	570	2	US-08-967-364-1	Sequence 1, Appl
33	37	44.0	570	2	US-08-967-364-7	Sequence 7, Appl
34	37	44.0	570	3	US-09-368-408-1	Sequence 1, Appl
35	37	44.0	570	3	US-09-368-408-7	Sequence 7, Appl
36	37	44.0	1912	4	US-08-913-832A-2	Sequence 2, Appl
37	37	44.0	1912	4	US-09-249-181A-2	Sequence 2, Appl
38	36	42.9	337	2	US-08-467-559B-2	Sequence 2, Appl
39	36	42.9	343	2	US-08-788-539A-2	Sequence 2, Appl
40	35	41.7	11	2	US-08-618-696-7	Sequence 7, Appl
41	35	41.7	11	3	US-09-033-753-7	Sequence 7, Appl
42	35	41.7	39	3	US-08-851-843A-23	Sequence 23, Appl
43	35	41.7	39	4	US-08-974-549A-213	Sequence 213, Appl
44	35	41.7	39	4	US-08-854-050-23	Sequence 23, Appl
45	35	41.7	39	4	US-09-430-323-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3
Query Match 67.9%; Score 57; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.00055;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 ERAIRQVEYFGDF 16
| | | | |
Db 4 EAKICHQIEYFGDF 18
RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

;; PRIOR FILING DATE: 1997-10-06
;; PRIOR APPLICATION NUMBER: 08/321,427
;; PRIOR FILING DATE: 1994-10-11
;; PRIOR APPLICATION NUMBER: 60/086,527
;; PRIOR FILING DATE: 1998-05-22
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
;; OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 67.9%; Score 57; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.00055;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIRQVEYYFGDF 16
| | | | | | | | | |
Db 4 EAKIQIEYYFGDF 18

RESULT 3
US-08-851-843A-26
; Sequence 26, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200

;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 26:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 39 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
US-08-851-843A-26
Query Match 67.9%; Score 57; DB 3; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.0013;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 IIROVEYYFGD 15
| | | | | | | | | |
Db 1 ILROVEYYFGD 11

RESULT 4
US-08-974-549A-216
; Sequence 216, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997

;;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 216:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-549A-216

Query Match 67.9%; Score 57; DB 4; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.0013;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 5 ILRQVEYYFGD 15
Db 1 ILRQVEYYFGD 11

RESULT 5
US-08-854-050-26
; Sequence 26, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536

;;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-854-050-26

Query Match 67.9%; Score 57; DB 4; Length 39;
Best Local Similarity 90.9%; Pred. No. 0.0013;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 5 ILRQVEYYFGD 15
Db 1 ILRQVEYYFGD 11

RESULT 6
US-09-430-323-26
; Sequence 26, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-0029300S

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 39 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-09-430-323-26

Query Match 57.9%; Score 57; DB 4; Length 39;

Best Local Similarity 90.9%; Pred. No. 0.0013; 0; Indels 0; Gaps 0;

Matches 10; Conservative 1; Mismatches 0;

QY 5 IIRQVEYFGD 15

Db 1 ILRQVEYFGD 11

RESULT 7

US-08-974-549A-214

Sequence 214, Application US/08974549A

Patent No. 6166178

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin B.

APPLICANT: Andrews, William H.

TITLE OF INVENTION: Human Telomerase Catalytic Subunit

NUMBER OF SEQUENCES: 727

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,549A

FILING DATE: 19-NOV-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/911,312

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/912,951

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/915,503

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17618

FILING DATE: 01-OCT-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17885

FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph Ted

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-0026100S

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 214:

SEQUENCE CHARACTERISTICS:

LENGTH: 38 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-974-549A-214

Query Match 59.5%; Score 50; DB 4; Length 38;

Best Local Similarity 75.0%; Pred. No. 0.023;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 IIRQVEYFGDF 16

Db 1 ICHQXEYFGDF 12

RESULT 8

US-08-851-843A-25

Sequence 25, Application US/08851843A

Patent No. 6093809

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin

APPLICANT: Andrews, William H.

TITLE OF INVENTION: No. 6093809el Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/851,843A

FILING DATE: 06-MAY-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

CLASSIFICATION:

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-851-843A-25

Query Match 57.1%; Score 48; DB 3; Length 38;
Best Local Similarity 72.7%; Pred. No. 0.053;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIRQVEYFFGD 15
Db 1 ICEQIEYFFGD 11

RESULT 9
US-08-974-549A-215
; Sequence 215, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050

```

```

; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0026100S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 215:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-549A-215

Query Match 57.1%; Score 48; DB 4; Length 38;
Best Local Similarity 72.7%; Pred. No. 0.053;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIRQVEYFFGD 15
Db 1 ICEQIEYFFGD 11

RESULT 10
US-08-854-050-25
; Sequence 25, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843

```

;
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/430,323
; FILING DATE: 29-OCT-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843

Query Match 57.1% Score 48; DB 4; Length 38;
Best Local Similarity 72.7%; Pred. No. 0.053; 2; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 IIRQVEYFGD 15
| | :|||||
Db 1 ICEQIEYFGD 11

RESULT 11
US-09-430-323-25
; Sequence 25, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430.323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843

;
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match 57.1% Score 48; DB 4; Length 38;
Best Local Similarity 72.7%; Pred. No. 0.053; 2; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 IIRQVEYFGD 15
| | :|||||
Db 1 ICEQIEYFGD 11

RESULT 12
US-08-851-843A-24
; Sequence 24, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643

```
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-851-843A-24

Query Match 51.2%; Score 43; DB 3; Length 37;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFGDF 16
DB 5 EYFGDF 11

RESULT 13
US-08-854-050-24
; Sequence 24, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-854-050-24

Query Match 51.2%; Score 43; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFGDF 16
DB 5 EYFGDF 11

RESULT 14
US-09-430-323-24
; Sequence 24, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
```

```

; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 37 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: not relevant
;   TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-430-323-24

Query Match          51.2%; Score 43; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFGDF 16
    |||||
Db 5 EYFGDF 11

RESULT 15
US-08-851-843A-27
; Sequence 27, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-08-851-843A-27

Query Match          50.0%; Score 42; DB 3; Length 38;
Best Local Similarity 54.5%; Pred. No. 0.63;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 IRQVEYFGDF 16
    :|||:|:|
Db 2 LKQVEYFSEF 12

Search completed: April 23, 2003, 13:36:36
Job time : 10.9775 secs

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:23 ; Search time 25.3483 Seconds
(without alignments)
84.109 Million cell updates/sec

Title: US-09-836-073-19

Perfect score: 84

Sequence: 1 QERAIRQVEYFGDF 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	78	92.9	390	22	ABB65316		Drosophila melanog
2	57	67.9	18	21	AA52200		Human la autoantig
3	57	67.9	92	21	AAG01351		Human secreted pro
4	57	67.9	408	17	AAW03716		Human autoantigen
5	57	67.9	439	22	ABG08417		Novel human diagno
6	57	67.9	460	21	AAB58987		Breast and ovarian
7	57	67.9	460	23	ABP41511		Human ovarian anti
8	54	64.3	481	14	AAR43394		La/SSB epitope 17.
9	48	57.1	483	21	AAG29675		Arabidopsis thalia
10	47	56.0	913	21	AAG47714		Arabidopsis thalia

11	47	56.0	923	21	AAG47713	Arabidopsis thalia
12	47	56.0	993	21	AAG47712	Arabidopsis thalia
13	47	56.0	1379	22	ABB68940	Drosophila melanog
14	41	48.8	324	22	ABB63243	Drosophila melanog
15	40.5	48.2	2703	16	AAW70236	P. falciparum Proj
16	40.5	48.2	2710	18	AAW22482	Plasmodium Proj3.
17	40.5	48.2	2710	21	AAW77904	P. falciparum Proj
18	40.5	48.2	3060	18	AAW22475	Plasmodium var-7.
19	40.5	48.2	3060	21	AAW77905	Plasmodium var-7 p
20	40	47.6	198	22	AAW25632	Human protein sequ
21	40	47.6	262	22	AAG1976	C glutamicum prote
22	40	47.6	373	23	ABB83256	Escherichia coli p
23	40	47.6	373	23	AAO17421	E. coli tyra protei
24	40	47.6	373	23	AAW49653	E. coli tyra prote
25	40	47.6	418	22	AAW39884	Human polypeptide,
26	40	47.6	457	20	AAW60555	Human normal bladd
27	40	47.6	503	22	AAG64905	Human granule memb
28	40	47.6	506	21	AAW20572	Antirrhinum majus
29	40	47.6	506	21	AAW20574	Perilla frutescens
30	40	47.6	506	22	AAW20175	Perilla hybrida fl
31	40	47.6	512	21	AAW20573	Torenia hybrida fl
32	40	47.6	556	23	ABP30047	Streptococcus poly
33	40	47.6	590	23	ABG66704	Human novel polype
34	40	47.6	642	21	AAW43122	Human ORFX ORF2886
35	40	47.6	662	23	ABG66750	Human novel polype
36	40	47.6	979	22	AAW66398	Receptor 222. Unl
37	40	47.6	1062	22	AAU01014	Human secreted sol
38	40	47.6	1062	22	AAW62238	Human calcium chan
39	40	47.6	1066	22	ABP37987	Human GS96663 prot
40	40	47.6	1082	22	AAU01015	Human secreted sol
41	40	47.6	1082	22	AAW62239	Human calcium chan
42	40	47.6	1109	22	AAU01016	Human secreted sol
43	40	47.6	1109	22	AAW62240	Human calcium chan
44	40	47.6	1115	22	AAU01025	Human wild type al
45	40	47.6	1115	22	AAW62249	Human calcium chan

ALIGNMENTS

RESULT 1
ABB65316
ID ABB65316 standard; Protein; 390 AA.
XX
AC ABB65316;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 22740.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR N-PSDB; ABL09419.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell

PT Interactions -
 PS Disclosure; SEQ ID NO 22740; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 390 AA;
 Query Match 92.9%; Score 78; DB 22; Length 390;
 Best Local Similarity 100.0%; Pred. No. 1.9e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QERAIIRQVEYFGD 15
 DB 50 QERAIIRQVEYFGD 64
 RESULT 2
 AAY52200
 ID AAY52200 standard; peptide; 18 AA.
 AC AAY52200;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Human la autoantigen peptide (LAP).
 XX
 KW La autoantigen; LAP; internal ribosome entry site; IRES; translation;
 KW viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
 KW coronavirus; hepatitis virus; rhadovirus; adenovirus; coxsackie virus;
 KW parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
 KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
 KW vesicular stomatitis virus.
 XX
 OS Homo sapiens.
 XX
 WN WO9961613-A2.
 XX
 PD 02-DEC-1999.
 XX
 PF 21-MAY-1999; 99WO-US11281.
 XX
 PR 22-MAY-1998; 98US-0086527.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Das S, Dasgupta A;
 XX
 DR WPI; 2000-062712/05.
 XX
 PT New yeast inhibitory peptide useful for inhibiting viral protein
 PT translation and replication -
 XX
 PS Claim 5; Page 57; 81pp; English.
 XX
 CC This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (IRNA see AA245200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA

CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rhabdovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.
 XX
 SQ Sequence 18 AA;
 Query Match 67.9%; Score 57; DB 21; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.0028;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ERAIRQVEYFGDF 16
 DB 4 EARICHOIEYFGDF 18
 RESULT 3
 AAG01351
 ID AAG01351 standard; Protein; 92 AA.
 XX
 AC AAG01351;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 5432.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 DR N-PSDB; AAC01357.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13; SEQ ID 5432; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and are therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SQ Sequence 92 AA;

Query Match 67.9%; Score 57; DB 21; Length 92;
 Best Local Similarity 66.7%; Pred. No. 0.018;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIRQVEYFGDF 16
 DB 14 EAKICHQIEYFGDF 28

RESULT 4
 AAW03716
 ID AAW03716 standard; protein; 408 AA.
 XX
 AC AAW03716;
 XX
 DT 12-MAR-1997 (first entry)
 XX
 DE Human autoantigen La(SS-B).
 XX
 KW Autoimmune disease; La autoantigen; Sjogren's syndrome;
 KW systemic lupus erythematosus; diagnosis.
 XX
 OS Homo sapiens.
 XX
 PN US5541291-A.
 XX
 PD 30-JUL-1996.
 XX
 PF 31-DEC-1984; 84US-0687908.
 XX
 PR 27-MAY-1987; 87US-0054871.
 PR 31-DEC-1984; 84US-0687908.
 XX
 PA (UYDU-) UNIV DUKE.
 XX
 PI Keene JD;
 XX
 DR WPI; 1996-362015/36.
 XX
 PT Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
 PT overlap syndrome - useful for diagnosis and treatment of autoimmune
 PT diseases
 XX
 PS Disclosure; Columns 15-16; 21pp; English.
 CC
 CC The human lupus antigen (La) is diagnostic for Sjogren's syndrome,
 CC as well as occurring in systemic lupus erythematosus patients. The
 CC La protein is clinically related to the Ro protein that is highly
 CC common among autoimmune patients. La and Ro antigens sometimes
 CC reside on the same cellular ribonucleoprotein particle; most La
 CC patients contain some Ro antibodies and vice versa. La cDNA has
 CC been isolated from a human liver library.
 XX
 SQ Sequence 408 AA;

Query Match 67.9%; Score 57; DB 17; Length 408;
 Best Local Similarity 66.7%; Pred. No. 0.1;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIRQVEYFGDF 16
 DB 14 EAKICHQIEYFGDF 28

RESULT 5
 ABG08417
 ID ABG08417 standard; Protein; 439 AA.
 XX
 AC ABG08417;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #8408.

XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS72604.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 38776; 103pp; English.
 CC
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 439 AA;

Query Match 67.9%; Score 57; DB 22; Length 439;
 Best Local Similarity 66.7%; Pred. No. 0.11;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIRQVEYFGDF 16
 DB 44 EAKICHQIEYFGDF 58

RESULT 6
 AAB58987
 ID AAB58987 standard; Protein; 460 AA.
 XX
 AC AAB58987;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 695.
 XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW

KW nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidabetic; antiinflammatory; antiulcer; vulnar; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease.
 XX
 OS Homo sapiens.
 XX WO200055173-A1.
 XX 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-US05881.
 XX
 XX 12-MAR-1999; 99US-0124270.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM;
 XX WPI; 2000-611515/58.
 XX N-PSDB; AAF21890.
 DR
 XX
 PT New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases -
 XX
 PS Claim 11; Page 1149-11150; 1299pp; English.
 XX
 CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 CC antidabetic; antiinflammatory; antiulcer; vulnar; anticonvulsant;
 CC antibacterial; antifungal; antiparasitic and cardiant activity. The
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and antagonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemias; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.
 XX
 SQ Sequence 460 AA;
 Query Match 67.9%; Score 57; DB 21; Length 460;
 Best Local Similarity 66.7%; Pred. No. 0.12;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 OY 2 ERAIRQVEYFGDF 16
 Db 66 EAKICHQIEYFGDF 80
 RESULT 7
 ID AAF41511 standard; Protein; 460 AA.
 XX
 AC AAF41511;
 XX
 XX 22-AUG-2002 (first entry)
 DT
 XX Human ovarian antigen HVAF56, SEQ ID NO:2643.
 DE
 XX

KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive.
 XX
 OS Homo sapiens.
 XX WO200200677-A1.
 XX 03-JAN-2002.
 XX
 XX 07-JUN-2001; 2001WO-US18569.
 XX
 XX 07-JUN-2000; 2000US-209467P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Birse CE, Rosen CA;
 XX WPI; 2002-147878/19.
 XX N-PSDB; ABQ54588.
 DR
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 11; SEQ ID NO 2643; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

Sequence 460 AA;

Query Match 67.9%; Score 57; DB 23; Length 460;

Best Local Similarity 66.7%; Pred. No. 0.12;

Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 ERAIRQVEYFGDF 16

| | | | |

Db 66 EAKICHQIEYFGDF 80

RESULT 8
AAR43394
ID AAR43394 standard; peptide; 21 AA.

AC AAR43394;
XX
DT 12-MAY-1994 (first entry)
XX
DE La/SSB epitope 17.

XX Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

XX OS Homo sapiens.

XX PN W09321223-A.

XX PD 28-OCT-1993.

XX PF 13-APR-1993; 93WO-US03484.

XX PR 13-APR-1992; 92US-0867819.

XX PA (OKLA) UNIV OKLAHOMA STATE.

XX PI Harley JB;

XX DR WPI; 1993-351658/44.

XX PT New linear epitope(s) for human auto-antibodies - from the
PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
PT for diagnosing and treating auto-immune disorders e.g. systemic
PT lupus erythematosus

XX PS Claim 1; Page 30; 43pp; English.

XX CC The sequences given in AAR43391-562 are linear epitopes which are
CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
CC polypeptide. These antigens are common in systemic lupus
CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
CC of proteins has been shown to have several molecular forms which are
CC defined by the molecular weight of the antigen identified. The major
CC form has a molecular weight of 60 kD and two additional forms have
CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
CC group of autoantibodies and binds small RNAs with a polypyridine
CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
CC phosphoprotein which associates with RNA polymerase III transcripts.
CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
CC U5 RNA. Anti-Sm antibodies may be directed against one or a
CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
CC used for preventing, treating or screening autoimmune disorders,
CC especially SLE or Sjogrens syndrome (SS). They bind to a human
CC autoantibody and may therefore be used as vaccines.

XX SQ Sequence 21 AA;

Query Match 64.3%; Score 54; DB 14; Length 21;
Best Local Similarity 75.0%; Pred. No. 0.011; 2; Indels 0; Gaps 0;
Matches 9; Conservative 1; Mismatches 1;

Qy 5 IIRQVEYFGDF 16
| :|||||
Db 1 ICHQIEYFGDF 12

RESULT 9

AAG29675

ID AAG29675 standard; Protein; 483 AA.

XX XX

AC AAG29675;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 35349.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126284.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

XX Arabidopsis thaliana protein fragment SEQ ID NO: 60168.
DE
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
PN
XX
PD
XX
PF 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132048.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138340.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

```

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 56.0%; Score 47; DB 21; Length 913;
Best Local Similarity 72.7%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 IIRQVEYVFSD 15
   | | | | |
Db 117 IVNQVEYVFSD 127

RESULT 11
AAG47713
ID AAG47713 standard; Protein; 923 AA.
XX AC AAG47713;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 60167.
XX DE
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX KW

```

```

OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.

```

```

PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 56.0%; Score 47; DB 21; Length 923;
Best Local Similarity 72.7%; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 IIRQVEYYFGD 15
Db 127 IVNQVEYYFSD 137

RESULT 12
AAG47712
ID AAG47712 standard; Protein; 993 AA.
XX
AC AAG47712;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 60166.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.

```

XX			
PR	25-FEB-1999;	99US-0121825.	99US-0142977.
PR	05-MAR-1999;	99US-0123180.	99US-0143542.
PR	09-MAR-1999;	99US-0123548.	99US-0143624.
PR	23-MAR-1999;	99US-0125788.	99US-0144003.
PR	25-MAR-1999;	99US-0126264.	99US-0144085.
PR	29-MAR-1999;	99US-0126785.	99US-0144086.
PR	01-APR-1999;	99US-0127462.	99US-0144325.
PR	06-APR-1999;	99US-0128234.	99US-0144331.
PR	08-APR-1999;	99US-0128714.	99US-0144332.
PR	16-APR-1999;	99US-0129845.	99US-0144333.
PR	19-APR-1999;	99US-0130077.	99US-0144334.
PR	21-APR-1999;	99US-0130449.	99US-0144352.
PR	23-APR-1999;	99US-0130510.	99US-0144632.
PR	28-APR-1999;	99US-0130891.	99US-0144884.
PR	30-APR-1999;	99US-0131449.	99US-0144884.
PR	30-APR-1999;	99US-0132048.	99US-0145086.
PR	04-MAY-1999;	99US-0132407.	99US-0145088.
PR	05-MAY-1999;	99US-0132484.	99US-0145085.
PR	06-MAY-1999;	99US-0132485.	99US-0145087.
PR	06-MAY-1999;	99US-0132486.	99US-0145089.
PR	07-MAY-1999;	99US-0132487.	99US-0145192.
PR	11-MAY-1999;	99US-0132863.	99US-0145145.
PR	14-MAY-1999;	99US-0134256.	99US-0145218.
PR	14-MAY-1999;	99US-0134218.	99US-0145224.
PR	14-MAY-1999;	99US-0134219.	99US-0145276.
PR	14-MAY-1999;	99US-0134221.	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	99US-0145951.
PR	20-MAY-1999;	99US-0135124.	99US-0146386.
PR	21-MAY-1999;	99US-0135353.	99US-0146388.
PR	24-MAY-1999;	99US-0135629.	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	99US-0147038.
PR	27-MAY-1999;	99US-0136392.	99US-0147204.
PR	28-MAY-1999;	99US-0136782.	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	99US-0147192.
PR	03-JUN-1999;	99US-0137528.	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	99US-0147416.
PR	08-JUN-1999;	99US-0138094.	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	99US-0148565.
PR	17-JUN-1999;	99US-0139459.	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	99US-0151080.
PR	23-JUN-1999;	99US-0140353.	99US-0151303.
PR	23-JUN-1999;	99US-0140354.	99US-0151338.
PR	24-JUN-1999;	99US-0140695.	99US-0151930.
PR	28-JUN-1999;	99US-0140823.	99US-0152363.
PR	29-JUN-1999;	99US-0140991.	99US-0153070.
PR	30-JUN-1999;	99US-0141287.	99US-0153758.
PR	01-JUL-1999;	99US-0141842.	99US-0154018.
PR	01-JUL-1999;	99US-0142154.	99US-0154039.
PR	02-JUL-1999;	99US-0142055.	99US-0154779.
PR	06-JUL-1999;	99US-0142390.	99US-0155139.
PR	08-JUL-1999;	99US-0142803.	99US-0155486.
PR	09-JUL-1999;	99US-0142920.	99US-0155659.
PR			99US-0156458.
PR	12-JUL-1999;		
PR	13-JUL-1999;		
PR	14-JUL-1999;		
PR	15-JUL-1999;		
PR	16-JUL-1999;		
PR	16-JUL-1999;		
PR	19-JUL-1999;		
PR	19-JUL-1999;		
PR	19-JUL-1999;		
PR	19-JUL-1999;		
PR	20-JUL-1999;		
PR	20-JUL-1999;		
PR	20-JUL-1999;		
PR	21-JUL-1999;		
PR	21-JUL-1999;		
PR	22-JUL-1999;		
PR	22-JUL-1999;		
PR	22-JUL-1999;		
PR	22-JUL-1999;		
PR	22-JUL-1999;		
PR	23-JUL-1999;		
PR	23-JUL-1999;		
PR	23-JUL-1999;		
PR	26-JUL-1999;		
PR	27-JUL-1999;		
PR	27-JUL-1999;		
PR	27-JUL-1999;		
PR	28-JUL-1999;		
PR	02-AUG-1999;		
PR	02-AUG-1999;		
PR	03-AUG-1999;		
PR	04-AUG-1999;		
PR	04-AUG-1999;		
PR	05-AUG-1999;		
PR	05-AUG-1999;		
PR	06-AUG-1999;		
PR	06-AUG-1999;		
PR	09-AUG-1999;		
PR	09-AUG-1999;		
PR	10-AUG-1999;		
PR	11-AUG-1999;		
PR	12-AUG-1999;		
PR	13-AUG-1999;		
PR	13-AUG-1999;		
PR	16-AUG-1999;		
PR	17-AUG-1999;		
PR	18-AUG-1999;		
PR	20-AUG-1999;		
PR	20-AUG-1999;		
PR	20-AUG-1999;		
PR	23-AUG-1999;		
PR	23-AUG-1999;		
PR	25-AUG-1999;		
PR	26-AUG-1999;		
PR	27-AUG-1999;		
PR	27-AUG-1999;		
PR	27-AUG-1999;		
PR	30-AUG-1999;		
PR	31-AUG-1999;		
PR	01-SEP-1999;		
PR	07-SEP-1999;		
PR	10-SEP-1999;		
PR	13-SEP-1999;		
PR	15-SEP-1999;		
PR	16-SEP-1999;		
PR	20-SEP-1999;		
PR	22-SEP-1999;		
PR	23-SEP-1999;		
PR	24-SEP-1999;		
PR	28-SEP-1999;		

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:34:48 ; Search time 17.7978 Seconds
(without alignments)
185.234 Million cell updates/sec

Title: US-09-836-073-19

Perfect score: 84
Sequence: 1 QERAIHQVEYFGDF 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 9297

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	41.7	16	2 Q9R963	Q9R963 helicobacte
2	32	38.1	11	4 Q9UC46	Q9UC46 homo sapien
3	31	36.9	20	10 Q9SANS	Q9SANS arabidopsis
4	28	33.3	17	7 Q19716	Q19716 homo sapien
5	25	29.8	25	6 Q9RTG0	Q9RTG0 atelies belz
6	25	29.8	25	8 Q9TGB8	Q9TGB8 alnus crisp
7	25	29.8	25	8 Q9TGB7	Q9TGB7 alnus glut
8	25	29.8	25	8 Q9TGB6	Q9TGB6 alnus marit
9	25	29.8	25	8 Q9TGB5	Q9TGB5 betula alle
10	25	29.8	25	8 Q9TGB4	Q9TGB4 betula glan
11	25	29.8	25	8 Q9TGB3	Q9TGB3 betula papy
12	25	29.8	25	8 Q9TGB2	Q9TGB2 betula veir
13	25	29.8	25	8 Q9TGB1	Q9TGB1 betula pube
14	25	29.8	25	8 Q9TGB0	Q9TGB0 corylus ave
15	25	29.8	25	8 Q9TGA9	Q9TGA9 corylus col
16	25	29.8	25	8 Q9TGA8	Q9TGA8 corylus cor

17	25	29.8	25	8 Q9TGA7	Q9TGA7 ostriya virg.
18	25	29.8	25	8 Q9TGA6	Q9TGA6 quercus rub
19	25	29.8	25	8 Q9TGB3	Q9TGB3 carpinus ca
20	24	28.6	13	8 Q9THR8	Q9THR8 bryopsis sp
21	24	28.6	20	10 Q9S8X5	Q9S8X5 glycine max
22	24	28.6	23	2 Q52706	Q52706 rhodobacter
23	24	28.6	24	4 Q13660	Q13660 homo sapien
24	24	28.6	25	2 Q92EW1	Q92EW1 enterobacte
25	24	28.6	25	2 Q9R4S5	Q9R4S5 streptomyce
26	24	28.6	25	8 Q9XJZ7	Q9XJZ7 aureocombra
27	23	27.4	8	2 Q09258	Q09258 synechococc
28	23	27.4	15	2 Q9R521	Q9R521 francisella
29	23	27.4	18	2 Q9S1G4	Q9S1G4 saccharopol
30	23	27.4	18	2 Q57323	Q57323 yersinia ps
31	23	27.4	18	12 Q9W9C1	Q9W9C1 human adeno
32	23	27.4	19	2 Q47078	Q47078 escherichia
33	23	27.4	23	11 Q9QV93	Q9QV93 rattus sp.
34	22	26.2	18	4 Q9H1I3	Q9H1I3 homo sapien
35	22	26.2	19	10 Q9S8R1	Q9S8R1 solanum tub
36	22	26.2	20	2 Q9R4L7	Q9R4L7 leuconostoc
37	22	26.2	20	5 Q9TWR0	Q9TWR0 blattella g
38	22	26.2	21	9 Q9T167	Q9T167 bacterioph
39	22	26.2	22	10 P82194	P82194 spinacia ol
40	22	26.2	22	15 Q9PXA8	Q9PXA8 human immu
41	22	26.2	24	2 Q9ZG39	Q9ZG39 chlamydia t
42	22	26.2	24	15 Q9PXA7	Q9PXA7 human immu
43	22	26.2	24	16 Q9K8M1	Q9K8M1 bacillus ha
44	22	26.2	25	10 Q40354	Q40354 medicago sa
45	22	26.2	25	16 P72429	P72429 salmonella

ALIGNMENTS

RESULT 1

Q9R963 ID Q9R963 PRELIMINARY; PRT; 16 AA.
AC Q9R963;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TremBLrel. 14, Last annotation update)
DE pepC (Fragment).
GN FEPC.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=F31;
RX MEDLINE=98453456; PubMed=9780260;
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,
RA Kuriyama M.;
RT "Full-length sequence analysis of the vacA gene from cytotoxic and
RT noncytotoxic Helicobacter pylori.";
RL J. Infect. Dis. 178:1391-1398(1998).
DR EMBL; AF049623; AAD04263.1;
FT NON_TER 1
FT NON_TER 16
SQ SEQUENCE 16 AA; 1938 MW; 3C33DA03AEE61428 CRC64;

Query Match 41.7%; Score 35; DB 2; Length 16;
Best Local Similarity 61.5%; Pred. No. 33;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 AIIHQVEYFGDF 16
| | | | | | | |
Db 3 AIIHQVEYAFNF 15

RESULT 2

Q9UC46 ID Q9UC46 PRELIMINARY; PRT; 11 AA.

AC OSUC46;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE Neutrophil inhibitor peptide, NIP-POLYMORPHONUCLEAR neutrophil
 DE inhibitor peptide.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96326114; PubMed=8703476;
 RA Cooper J.A.Jr., Culbreth R.R.;
 RT "Characterization of a neutrophil inhibitor peptide harvested from
 RT human bronchial lavage: homology to influenza A nucleoprotein.";
 RL Am. J. Respir. Cell Mol. Biol. 15:207-215(1996).
 SQ SEQUENCE 11 AA; 1262 MW; 951A1C3279C9DB45 CRC64;
 Query Match 38.1%; Score 32; DB 4; Length 11;
 Best Local Similarity 55.6%; Pred. No. 72;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 7 ROVEYVFGD 15
 I: |:|:|
 Db 1 REGSYFGD 9
 RESULT 3
 Q9SANS PRELIMINARY; PRT; 20 AA.
 AC Q9SANS;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE Mitochondrial LMW heat shock protein (Fragment).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97369374; PubMed=9225862;
 RA Visoli G., Maestri E., Marmioli N.;
 RT "Differential display-mediated isolation of a genomic sequence for a
 RT putative mitochondrial LMW HSP specifically expressed in condition of
 RT induced thermotolerance in Arabidopsis thaliana (L.) heynh.";
 RL Plant Mol. Biol. 34:517-527(1997).
 DR EMBL: Y11865; CAA72614.1; --
 KW Heat shock.
 FT NON_TER 1 1
 SQ SEQUENCE 20 AA; 2439 MW; 0DF3276AD6ADEAAB CRC64;
 Query Match 36.9%; Score 31; DB 10; Length 20;
 Best Local Similarity 60.0%; Pred. No. 2e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 QERAIIRQVE 10
 I: |:|:|
 Db 9 QERNDVROIE 18
 RESULT 4
 O19716 PRELIMINARY; PRT; 17 AA.
 ID O19716;
 AC O19716;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Human histocompatibility system hla-dr heavy chain (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=82197531; PubMed=6952207;
 RA Lee J.S., Trowsdale J., Bodmer W.F.;
 RT "cDNA clones coding for the heavy chain of human hla-dr antigen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:545-549(1982).
 DR EMBL: J00193; AAA36272.1; --
 FT NON_TER 1 1
 FT NON_TER 17 17
 SQ SEQUENCE 17 AA; 2058 MW; 4874E328EE648F54 CRC64;
 Query Match 33.3%; Score 28; DB 7; Length 17;
 Best Local Similarity 41.7%; Pred. No. 5.4e+02;
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 1 QERAIIRQVEY 12
 I: |:|:|
 Db 2 KEEHVIIQAEFY 13
 RESULT 5
 Q9TTGO PRELIMINARY; PRT; 25 AA.
 ID Q9TTGO;
 AC Q9TTGO;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Alpha-cardiac actin (Fragment).
 GN ACTC.
 OS Ateles belzebuth chamek (Chamek spider monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae; Ateles.
 OX NCBI_TaxID=118643;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20169179; PubMed=10702662;
 RA Canavez F.C., Moreira M.A., Bonvicino C.R., Parham P., Seunarez H.N.;
 RT "Evolutionary disruptions of human syntenic groups 3, 12, 14, and 15
 RT in Ateles belzebuth chamek (Platyrrhini, Primates).";
 RL Cytogenet. Cell Genet. 87:182-188(1999).
 DR EMBL: AF099178; AAF21852.1; --
 DR InterPro: IPR004000; Actin_like.
 DR Pfam: PF00022; actin; 1.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 2982 MW; 5242AF35E2169DCB CRC64;
 Query Match 29.8%; Score 25; DB 6; Length 25;
 Best Local Similarity 44.4%; Pred. No. 2.6e+03;
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 2 ERAIRQVE 10
 I: |:|:|
 Db 15 EREIVRDIK 23
 RESULT 6
 Q9TGB8 PRELIMINARY; PRT; 25 AA.
 ID Q9TGB8;
 AC Q9TGB8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Ribosomal protein small 3 (Fragment).
 GN RPS3.
 OS Alnus crispa.
 OC Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fagales; Betulaceae; Alnus.
 OX NCBI_TaxID=3518;

```

RN SEQUENCE FROM N.A.
RX MEDLINE-99261656; PubMed-10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080075; AAD50062.1; -
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 29.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFYG 14
Db 4 DYYG 8

RESULT 7
Q9TGB7
ID Q9TGB7 PRELIMINARY; PRT; 25 AA.
AC Q9TGB7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Alnus glutinosa (Alder).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=3517;
RN SEQUENCE FROM N.A.
RP MEDLINE-99261656; PubMed-10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080076; AAD50063.1; -
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 29.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFYG 14
Db 4 DYYG 8

RESULT 8
Q9TGB6
ID Q9TGB6 PRELIMINARY; PRT; 25 AA.
AC Q9TGB6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Alnus maritima.

```

```

OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=21015;
RN SEQUENCE FROM N.A.
RP MEDLINE-99261656; PubMed-10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080077; AAD50064.1; -
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 29.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFYG 14
Db 4 DYYG 8

RESULT 9
Q9TGB5
ID Q9TGB5 PRELIMINARY; PRT; 25 AA.
AC Q9TGB5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Betula alleghaniensis.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=21017;
RN SEQUENCE FROM N.A.
RP MEDLINE-99261656; PubMed-10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080078; AAD50065.1; -
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 29.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFYG 14
Db 4 DYYG 8

RESULT 10
Q9TGB4
ID Q9TGB4 PRELIMINARY; PRT; 25 AA.
AC Q9TGB4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment)
GN RPS3.
OS Betula glandulosa.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=21018;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RT angiosperms and homology to nad5 intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080079; AAD50066.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
DR Mitochondrion.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 29.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFG 14
DB 4 DYIG 8

RESULT 11
Q9TGB3 ID Q9TGB3 PRELIMINARY; PRT; 25 AA.
AC Q9TGB3;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment)
GN RPS3.
OS Betula papyrifera (Paper birch).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=3507;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RT angiosperms and homology to nad5 intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080080; AAD50067.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
DR Mitochondrion.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 29.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFG 14
DB 4 DYIG 8

RESULT 12
Q9TGB2 ID Q9TGB2 PRELIMINARY; PRT; 25 AA.
AC Q9TGB2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment)
GN RPS3.
OS Betula verrucosa (White birch) (Betula pendula).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=3505;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RT angiosperms and homology to nad5 intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080081; AAD50068.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
DR Mitochondrion.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 29.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFG 14
DB 4 DYIG 8

RESULT 13
Q9TGB1 ID Q9TGB1 PRELIMINARY; PRT; 25 AA.
AC Q9TGB1;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment)
GN RPS3.
OS Betula pubescens (downy birch).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=38787;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RT angiosperms and homology to nad5 intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080082; AAD50069.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
DR Mitochondrion.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 29.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFG 14
DB 4 DYIG 8

RESULT 14
Q9TGB4 ID Q9TGB4 PRELIMINARY; PRT; 25 AA.
AC Q9TGB4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment)
GN RPS3.
OS Betula pubescens (downy birch).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=38787;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RT angiosperms and homology to nad5 intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080083; AAD50070.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
DR Mitochondrion.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 29.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFG 14
DB 4 DYIG 8

```

QY 10 EYFG 14
:||:|
Db 4 DYYG 8

RESULT 14

Q9TGB0 PRELIMINARY; PRT; 25 AA.
AC Q9TGB0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Corylus avellana (European hazel).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Corylus.
OX NCBI_TaxID=13451;
RN [1]
R1 SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RT angiosperms and homology to nad5 intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080084; AAD50071.1; -;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 29.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFG 14
:||:|
Db 4 DYYG 8

RESULT 15

Q9TGA9 PRELIMINARY; PRT; 25 AA.
AC Q9TGA9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Corylus colurna.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Corylus.
OX NCBI_TaxID=101193;
RN [1]
R1 SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RT angiosperms and homology to nad5 intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080085; AAD50072.1; -;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 29.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.6e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 EYFG 14
:||:|
Db 4 DYYG 8

Search completed: April 23, 2003, 13:47:19
Job time : 18.7978 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:33:13 ; Search time 4.13483 Seconds
(without alignments)
160.495 Million cell updates/sec

Title: US-09-836-073-19
Perfect score: 84
Sequence: 1 QERAIIRQVEYFQDF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1520

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	27	32.1	19	1 HBB2_UROHA	P18992 uromastix h
2	24	28.6	16	1 MLB_SQUAC	P01207 squalus aca
3	24	28.6	25	1 SNBP_RAT	P80968 rattus norv
4	23	27.4	18	1 MLB_SCYCA	P01206 scyllorhinu
5	22	26.2	19	1 NUO6_SOLTA	P80729 solanum tub
6	22	26.2	20	1 THIO_CANFA	P99505 canis famil
7	22	26.2	25	1 ANDT_ANDAU	P56684 androctonus
8	21	25.0	18	1 SPAH_HELAN	P81098 helianthus
9	21	25.0	20	1 COG4_CHIOP	P34156 chionoecete
10	21	25.0	23	1 NUO5_SOLTA	P80262 solanum tub
11	20	23.8	15	1 LPF_ECOLI	P03057 escherichia
12	20	23.8	24	1 DHE3_PYRWO	Q09115 pyrococcus
13	20	23.8	25	1 KSP2_THETH	P21778 thermus the
14	19	22.6	8	1 ALI1_CARMA	P81820 carcinus ma
15	19	22.6	9	1 ALI1_CARMA	P81814 carcinus ma
16	19	22.6	16	1 AU21_LITRA	P82388 litoria ran
17	19	22.6	16	1 AU25_LITRA	P82392 litoria ran
18	19	22.6	19	1 FIBB_VULVU	P14482 vulpes vulp
19	19	22.6	20	1 FIBB_FELCA	P14469 felis silve
20	19	22.6	23	1 PRO3_DACGL	P18690 dactylis gl
21	19	22.6	24	1 VORC_METTM	P80909 methanobact
22	19	22.6	25	1 NEUW_PIG	P34964 sus scrofa
23	18	21.4	4	1 OCPI_OCTWI	P58648 octopus min
24	18	21.4	8	1 LMT2_LOCMI	P22396 locusta mig
25	18	21.4	13	1 NF1_LYMST	P80178 lymanaea sta
26	18	21.4	13	1 NF2_LYMST	P80179 lymanaea sta
27	18	21.4	13	1 NP3_LYMST	P80180 lymanaea sta
28	18	21.4	13	1 NP4_LYMST	P80181 lymanaea sta
29	18	21.4	13	1 NP5_LYMST	P80182 lymanaea sta
30	18	21.4	16	1 FIBA_CERST	P14535 ceratotheri
31	18	21.4	16	1 FIBA_MACFU	P12803 macaca fusc
32	18	21.4	16	1 FIBA_MANLE	P14455 mandrillus
33	18	21.4	18	1 MCRB_METTE	P22949 methanosarc

34 18 21.4 18 1 OBP_LYMDI P34173 lymantria d
35 18 21.4 19 1 FIBA_BISBO P14441 bison bonas
36 18 21.4 19 1 LCRP_PETMA Q10996 petromyzon
37 18 21.4 19 1 PA2S_HUMAN P24606 homo sapien
38 18 21.4 19 1 RL10_CITFR P43448 citrobacter
39 18 21.4 20 1 COXN_THUOB P80980 thunnus obe
40 18 21.4 20 1 PSAL_SYNVU P25937 synchococc
41 18 21.4 20 1 RL10_PROVU P51411 proteus vul
42 18 21.4 20 1 YOAH_KLEAE P56506 klebsiella
43 18 21.4 22 1 RL10_KLEPN P61190 klebsiella
44 18 21.4 23 1 CYSF_TRIFO P33403 tritrichomo
45 18 21.4 23 1 RL10_ENTCL Q47608 enterobacte

ALIGNMENTS

RESULT 1
HBB2_UROHA STANDARD; PRT; 19 AA.
ID HBB2_UROHA
AC P18992;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hemoglobin beta-2 chain (Fragment)
OS Uromastix hardwickii (Indian spiny-tailed lizard).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Uromastycinae;
OC Uromastix.
OX NCBI_TaxID=40250;
RN [1]
RP MEDLINE=84029159; PubMed=6628672;
RA Nagvi S., Zaidi Z.H., von Bahr-Lindstroem H., Carlquist M.,
RA Joernvall H.;
RT "Characterization of hemoglobin from the lizard Uromastix
hardwickii".
RL FEBS Lett. 162:290-295(1983).
CC -!- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CC -!- VARIOUS PERIPHERAL TISSUES.
CC -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -!- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -!- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR; A05305; A05305.
DR InterPro; IPR000971; Globin.
DR PROSITE; PS01033; GLOBIN; PARTIAL.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT NON_TER 1 19
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 1914 MW; C40AD8EA30019057 CRC64;

Query Match 32.1%; Score 27; DB 1; Length 19;
Best Local Similarity 80.0%; Pred. No. 98;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 YFGDF 16
:||||
Db 1 PFGDF 5

RESULT 2
MLB_SQUAC STANDARD; PRT; 16 AA.
ID MLB_SQUAC
AC P01207;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanotropin beta.
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]


```

RP SEQUENCE.
RX MEDLINE=75127390; PubMed=4375978;
RA Bennett H.P.J., Lowry P.J., McMartin C., Scott A.P.;
RT "Structural studies of alpha-melanocyte-stimulating hormone and a
RT novel beta-melanocyte-stimulating hormone from the neurointermediate
RL lobe of the pituitary of the dogfish Squalus acanthias.";
RL Biochem. J. 141:439-444(1974).
CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
DR PIR: A01471; MTFBFS.
KW Hormone.
SQ SEQUENCE 16 AA; 1939 MW; 993AF43C8A640A0E CRC64;

Query Match 28.6%; Score 24; DB 1; Length 16;
Best Local Similarity 57.1%; Pred. No. 2.8e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 10 EYVFGDF 16
DB 4 DYKFGHF 10

RESULT 3
SMBP_RAT
ID SMBP_RAT STANDARD; PRT; 25 AA.
AC P80968;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SM-11044 binding protein (Fragments).
OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RX MEDLINE=97407910; PubMed=9261134;
RA Sugawara T., Matsuzaki-Fujita M., Guillaume J.-L., Camoin L.,
RA Morooka S., Strosberg A.D.;
RT "Characterization of a novel iodoctanopindolol and SM-11044 binding
RT protein, which may mediate relaxation of depolarized rat colon
RT tonus.";
RL J. Biol. Chem. 272:21244-21252(1997).
CC -1- FUNCTION: MAY MEDIATE RELAXATION OF DEPOLARIZED COLON TONUS.
CC IT BINDS IODOCTANOPINDOLOL AND SM-11044.
CC NON-MISCELLANEOUS; THE ORDER OF THE FRAGMENTS IS UNKNOWN.
FT NON_TER 1 1
FT UNSURE 6 6 OR Y.
FT NON_CONS 18 19
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3177 MW; D14F0CB9B778C2CB CRC64;

Query Match 28.6%; Score 24; DB 1; Length 25;
Best Local Similarity 40.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 RQVEYFGDF 16
DB 12 RYQFYFPXF 21

RESULT 4
MLB_SCYCA
ID MLB_SCYCA STANDARD; PRT; 18 AA.
AC P01206;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanotropin beta (Beta-MSH).
OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
CC Elasmobranchii; Galeomorphii; Galeoidea; Carchariniiformes;
CC Scyliorhinidae; Scyliorhinus.

```

```

OX NCBI_TaxID=7830;
RN [1]
RP SEQUENCE.
RX MEDLINE=75113445; PubMed=4452470;
RA Love R.M., Pickering B.T.;
RT "A beta-MSH in the pituitary gland of the spotted dogfish
RT (Scyliorhinus canicula): Isolation and structure.";
RL Gen. Comp. Endocrinol. 24:398-404(1974).
CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
DR PIR: A01470; MTFBFC.
KW Hormone.
SQ SEQUENCE 18 AA; 2195 MW; 353C93A30C314DF4 CRC64;

Query Match 27.4%; Score 23; DB 1; Length 18;
Best Local Similarity 33.3%; Pred. No. 4.8e+02;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 8 QVEYFGDF 16
DB 2 ZIBYKMGHF 10

RESULT 5
NUO6_SOLTU
ID NUO6_SOLTU STANDARD; PRT; 19 AA.
AC P80729;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase 11 kDa subunit (EC 1.6.5.3)
DE (EC 1.6.99.3) (Complex I-11KD) (CI-11KD) (Fragment).
OS Solanum tuberosum (Potato).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
CC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RX STRAIN=cv. Bintje; TISSUE=Tuber;
RA Herz U., Grohmann L.;
RL Submitted (DBC-1996) to the SWISS-PROT data bank.
CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
CC -1- SUBCELLULAR LOCATION: MATRIX SIDE OF THE MITOCHONDRIAL INNER
CC MEMBRANE.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
FT NON_TER 19
SQ SEQUENCE 19 AA; 2310 MW; 0DCACEF407D79F49 CRC64;

Query Match 26.2%; Score 22; DB 1; Length 19;
Best Local Similarity 44.4%; Pred. No. 7.6e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ERAIRQVE 10
DB 8 ENLILRRME 16

RESULT 6
THIO_CANFA
ID THIO_CANFA STANDARD; PRT; 20 AA.
AC P99505;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thioredoxin (Fragment).
GN TXN.
OS Canis familiaris (Dog).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Heart;
 RA Dunn M.J., Wheeler C.H.;
 RL Submitted (Aug-1997) to the SWISS-PROT data bank.
 CC -1- FUNCTION: Participates in various redox reactions through the
 CC reversible oxidation of its active center dithiol to a disulfide
 CC and catalyzes dithiol-disulfide exchange reactions.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY.
 CC HSP; P10599; IERW.
 DR HSC-2DPAGE; P99505; DOG.
 DR InterPro: IPR000063; ThioRed.
 DR PROSITE: PS00194; THIOREDOXIN; PARTIAL.
 KW Redox-active center; Electron transport.
 FT UNSURE 6
 FT NON_TER 20
 SQ SEQUENCE 20 AA; 2287 MW; A06991862EB1B6A6 CRC64;

Query Match 26.2%; Score 22; DB 1; Length 20;
 Best Local Similarity 25.0%; Pred. No. 8.1e+02;
 Matches 2; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 6 IRQVEYF 13
 : : : : :
 DB 1 VKQIEFKY 8

RESULT 7
 ANDT ANDAU
 ID ANDT ANDAU STANDARD; PRT; 25 AA.
 AC P56684; P81616;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Androctonus.
 OS Androctonus australis hector (Sahara scorpion).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Buthidae; Buthidae; Androctonus.
 OX NCBI_TaxID=70175;
 RN [1]
 RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
 RC TISSUE=Hemolymph;
 RA Ehret-Sabatier L., Loew D., Goyffon M., Fehlbaum P., Hoffmann J.A.,
 van Dorselaer A., Bulet P.;
 RT "Characterization of novel cysteine-rich antimicrobial peptides from
 RT scorpion blood.";
 RL J. Biol. Chem. 271:29537-29544(1996).
 RN [2]
 RP SYNTHESIS OF D-AMINO ACID ENANTIOMER, AND CHARACTERIZATION.
 RA MEDLINE=20115101; PubMed=10642525;
 RX Hetru C., Letellier L., Oren Z., Hoffmann J.A., Shai Y.;
 RT "Androctonin, a hydrophilic disulphide-bridged non-haemolytic
 RT anti-microbial peptide: a plausible mode of action.";
 RL Biochem. J. 345:653-664(2000).
 RN [3]
 RP STRUCTURE BY NMR.
 RX MEDLINE=20025109; PubMed=10563585;
 RA Mandard N., Sy D., Maufrais C., Bonmatin J.M., Bulet P., Hetru C.,
 Vovelle F.;
 RA "Androctonin, a novel antimicrobial peptide from scorpion Androctonus
 RT australis: solution structure and molecular dynamics simulations in
 RT the presence of a lipid monolayer.";
 RL J. Biomol. Struct. Dyn. 17:367-380(1999).
 CC -1- FUNCTION: ACTIVE AGAINST BOTH BACTERIA (GRAM-POSITIVE AND GRAM-
 CC NEGATIVE) AND FILAMENTOUS FUNGI. ACTS ON THE MEMBRANE OF THE
 CC BACTERIAL CELLS. IT DESTABILIZE A MEMBRANE BY MODIFYING ITS
 CC PROPERTIES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MASS SPECTROMETRY: MW=3076.7; METHOD=Electrospray.

DR PDB; 1C26; 12-JAN-00.
 KW Antibiotic; Fungicide; 3D-structure.
 FT DISULFID 4 20
 FT DISULFID 10 16
 SQ SEQUENCE 25 AA; 3081 MW; D4183D6EA5AA50AA CRC64;

Query Match 26.2%; Score 22; DB 1; Length 25;
 Best Local Similarity 37.5%; Pred. No. 1e+03;
 Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 3 RAIIRQVE 10
 : : : : :
 DB 1 RSVCRQIK 8

RESULT 8
 SFAH HELAN
 ID SFAH HELAN STANDARD; PRT; 18 AA.
 AC P81098; 1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Seed fatty acyl-ester hydrolase (EC 3.1.1.1) (Fragment).
 DE Helianthus annuus (Common sunflower).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 OC Heliantheae; Helianthus.
 OX NCBI_TaxID=4232;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CV. RUSTICA / VAR. EUROFLOR; TISSUE=Seed;
 RA Beissner F., Gardies A.-M., Teissere M., Ferte N., Noat G.;
 RT "An esterase neosynthesized in post-germinated sunflower seeds is
 RT related to a new family of lipolytic enzymes.";
 RL Plant Physiol. Biochem. 35:761-765(1997).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=95210327; PubMed=7696323;
 RA Teissere M., Borel M., Caillol B., Nari J., Gardies A.-M., Noat G.;
 RT "Purification and characterization of a fatty acyl-ester hydrolase
 RT from post-germinated sunflower seeds.";
 RL Biochim. Biophys. Acta 1255:105-112(1995).
 CC -1- FUNCTION: IMPLICATED IN THE BREAKDOWN OF OIL BODY-STORED LIPIDS
 CC DURING POST-GERMINATION.
 CC -1- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O = an alcohol + a
 CC carboxylic anion.
 CC -1- TISSUE SPECIFICITY: SEED.
 CC -1- DEVELOPMENTAL STAGE: POST-GERMINATION.
 CC -1- SIMILARITY: BELONGS TO THE "GDSL" FAMILY OF LIPOLYTIC ENZYMES.
 DR InterPro: IPR001087; Lipase_GDSL.
 DR PROSITE: PS01098; LIPASE_GDSL_SER; PARTIAL.
 KW Hydrolase; Lipid degradation; Glycoprotein.
 FT ACT_SITE 13 13
 FT NON_TER 18 18
 SQ SEQUENCE 18 AA; 2057 MW; 3E9A80EB6548E862 CRC64;

Query Match 25.0%; Score 21; DB 1; Length 18;
 Best Local Similarity 60.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

OY 8 QVEYF-FGD 15
 : : : : :
 DB 3 QVPXYFIFGD 12

RESULT 9
 COG4_CHIOP
 ID COG4_CHIOP STANDARD; PRT; 20 AA.
 AC P34156;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Collagenolytic protease 23 kDa (EC 3.4.24.7) (Fragment).
 OS Chionoecetes opilio (Crab-beetle).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
 CC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
 CC Brachyura; Eubrachyura; Majoidae; Majidae; Chionoecetes.
 OX NCBI_TaxID=41210;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Hepatopancreas;
 RX MEDLINE=92120073; PubMed=1663026;
 RA Klimova O.A., Vedishcheva Y.V., Strongin A.Y.;
 RT "Isolation and characteristics of collagenolytic enzymes from the
 RT hepatopancreas of the crab Chionoecetes opilio.";
 RL Dokl. Akad. Nauk SSSR 317:482-484(1991).
 CC -|- FUNCTION: THIS ENZYME IS A METAL PROTEASE CAPABLE OF DEGRADING
 CC THE NATIVE TRIPLE HELIX OF COLLAGEN.
 CC -|- CATALYTIC ACTIVITY: Cleaves preferentially one bond in native
 CC collagen. Cleavage of the triple helix of collagen at about three-
 CC quarters of the length of the molecule from the N-terminus, at
 CC 775-Gly-Ile-776 in the alpha-1(I) chain. Cleaves synthetic
 CC substrates and alpha-macroglobulins at bonds where P1' is a
 CC hydrophobic residue.
 CC -|- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A.
 DR MEROPS: M12.001;
 DR InterPro: IPR000130; Zn_MTPeptidse.
 DR PROSITE: PS00142; ZINC_PROTEASE; PARTIAL.
 KW Hydrolase; Metalloprotease; Zinc; Collagen degradation.
 FT NON_TER 20
 SQ SEQUENCE 20 AA: 2108 MW; 2BC7A93D022A97D8 CRC64;
 Query Match 25.0%; Score 21; DB 1; Length 20;
 Best Local Similarity 66.7%; Pred. No. 1.2e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 9 VEYFEG 14
 Db | | | |
 15 VPVYFG 20
 RESULT 10
 NU05_SOLITU STANDARD; PRT; 23 AA.
 ID P80262;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE NADH-ubiquinone oxidoreductase 29 kDa subunit (EC 1.6.5.3)
 DE (EC 1.6.99.3) (Complex I-29KD) (CI-29KD) (Fragment).
 OS Solanum tuberosum (Potato).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=cv. Bintje; TISSUE=Tuber;
 RX MEDLINE=94124587; PubMed=8294484;
 RA Herz U., Schroeder W., Liddell A., Leaver C.J., Brennicke A.,
 RA Grolmann L.;
 RT "Purification of the NADH:ubiquinone oxidoreductase (complex I) of
 RT the respiratory chain from the inner mitochondrial membrane of
 RT Solanum tuberosum.";
 RL J. Biol. Chem. 269:2263-2269(1994).
 CC -|- FUNCTION: THE IMMEDIATE ELECTRON ACCEPTOR FOR THE RESPIRATORY
 CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
 CC TO BE UBIQUINONE.
 CC -|- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -|- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
 CC -|- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
 CC -|- SUBCELLULAR LOCATION: MATRIX SIDE OF THE MITOCHONDRIAL INNER
 CC MEMBRANE.
 PIR: C49732; C49732.

KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 FT NON_TER 23
 SQ SEQUENCE 23 AA: 2653 MW; 4B41B8FF83412F58 CRC64;
 Query Match 25.0%; Score 21; DB 1; Length 23;
 Best Local Similarity 40.0%; Pred. No. 1.4e+03;
 Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 10 EYFEG 14
 Db | | | |
 18 DYFYG 22
 RESULT 11
 LPF_ECOLI STANDARD; PRT; 15 AA.
 ID LPF_ECOLI
 AC P03057;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phe leader peptide (Attenuator peptide).
 DE PHE leader peptide (Attenuator peptide).
 GN PHEL OR PHEAE OR B2598.
 OS Escherichia coli.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=79033820; PubMed=360214;
 RA Zurawski G., Brown K., Killingly D., Yanofsky C.;
 RT "Nucleotide sequence of the leader region of the phenylalanine operon
 RT of Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 75:4271-4275(1978).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91072346; PubMed=2254312;
 RA Gavini N., Davidson B.E.;
 RT "phea mutants of Escherichia coli have a defective pheA attenuator.";
 RL J. Biol. Chem. 265:21532-21535(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 CC -|- FUNCTION: THIS PROTEIN IS INVOLVED IN CONTROL OF THE BIOSYNTHESIS
 CC OF PHENYLALANINE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: V00314; CAA23600.1;
 DR EMBL: M10431; AAA24329.1;
 DR EMBL: M58024; AAA62783.1;
 DR EMBL: AE000346; AAC75647.1;
 DR PIR: A03593; LFECF.
 DR PIR: B36494; B36494.
 DR Ecogene: EG11271; pheL.
 KW Leader peptide; Complete proteome.
 SQ SEQUENCE 15 AA: 1924 MW; CFE14AE3BFF935E0 CRC64;
 Query Match 23.8%; Score 20; DB 1; Length 15;
 Best Local Similarity 18.2%; Pred. No. 1.3e+03;
 Matches 2; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Dave H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.; isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.;
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 9 9 AMIDATION.
SQ SEQUENCE 9 AA; 927 MW; 832D79CDCB46D861 CRC64;

Query Match 22.6%; Score 19; DB 1; Length 9;
Best Local Similarity 60.0%; Pred. No. 1.1e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 EYVFG 14
: | | |
Db 4 QYAFG 8

Search completed: April 23, 2003, 13:43:54
Job time : 5.23483 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:36:43 ; Search time 8.35955 Seconds
(without alignments)
183.999 Million cell updates/sec

Title: US-09-836-073-19
Perfect score: 84
Sequence: 1 QERAIHQVEYFVGDF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 4984

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73: *
1: pirl: *
2: pirl2: *
3: pirl3: *
4: pirl4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	35.7	25	2 C57001	endo-1,4-beta-xyla
2	30	35.7	25	2 A60286	heat-stable serine
3	29	34.5	18	2 PH1368	Ig heavy chain DJ
4	28	33.3	23	2 PH1725	Ig heavy chain V r
5	28	33.3	24	2 PH1696	Ig heavy chain V r
6	27	32.1	12	2 PH0771	T-cell receptor be
7	27	32.1	19	2 A05305	hemoglobin beta-2
8	27	32.1	19	2 B53145	high conductance c
9	27	32.1	21	2 S46550	actin-related prot
10	26	31.0	16	2 E53284	T-cell receptor be
11	26	31.0	22	2 PH1325	Ig heavy chain DJ
12	25	29.8	15	2 PH1366	Ig heavy chain DJ
13	25	29.8	17	2 S57519	T cell receptor be
14	25	29.8	17	2 S57556	T cell receptor be
15	25	29.8	17	2 A61211	anantin - Streptom
16	25	29.8	21	2 PH1730	Ig heavy chain V r
17	25	29.8	22	2 PH1678	Ig heavy chain V r
18	25	29.8	22	2 PH1679	Ig heavy chain V r
19	25	29.8	23	2 PH1364	Ig heavy chain DJ
20	25	29.8	23	2 PH1681	Ig heavy chain V r
21	25	29.8	23	2 PH1682	Ig heavy chain V r
22	25	29.8	23	2 PH1694	Ig heavy chain V r
23	25	29.8	23	2 PH1707	Ig heavy chain V r
24	25	29.8	23	2 PH1722	Ig heavy chain V r
25	25	29.8	23	2 PH1724	Ig heavy chain V r
26	25	29.8	23	2 PH1727	Ig heavy chain V r
27	25	29.8	23	2 PH1723	Ig heavy chain V r
28	25	29.8	24	2 PH1683	Ig heavy chain V r
29	25	29.8	24	2 PH1685	Ig heavy chain V r

30	25	29.8	24	2 PH1698	Ig heavy chain V r
31	25	29.8	24	2 PH1710	Ig heavy chain V r
32	25	29.8	25	2 S29283	hydrogenase (EC 1.
33	25	29.8	25	2 PH1686	Ig heavy chain V r
34	25	29.8	25	2 PH1700	Ig heavy chain V r
35	25	29.8	25	2 PH1716	Ig heavy chain V r
36	25	29.8	25	2 PH1734	Ig heavy chain V r
37	24	28.6	12	2 S57570	T cell receptor V-
38	24	28.6	14	2 S57569	T cell receptor V-
39	24	28.6	14	2 S57638	T-cell receptor be
40	24	28.6	15	2 PH0760	melanotropin beta
41	24	28.6	16	1 MTDFBS	T-cell receptor be
42	24	28.6	18	2 B40741	chaperonin, 10K -
43	24	28.6	23	2 B39313	interleukin-1 beta
44	24	28.6	23	4 A58505	Ig heavy chain V r
45	24	28.6	24	2 PH1712	Ig heavy chain V r

ALIGNMENTS

RESULT 1

C57001
endo-1,4-beta-xylanase (EC 3.2.1.8) 4 - Streptomyces roseiscleroticus (fragment)
C:Species: Streptomyces roseiscleroticus
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 23-May-1997
C:Accession: C57001
R:Grabski, A.C.; Forrester, I.T.; Patel, R.; Jeffries, T.W.
A:Title: Characterization and N-terminal amino acid sequences of beta-(1-4)endoxyla
A:Reference number: A57001; MUID:93229899; PMID:8471845
A:Accession: C57001
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-25 <GRA>
A:Experimental source: strain NRRL B-11019
A:Note: sequence extracted from NCBI backbone (NCBIP:130009)
C:Function:
A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
A:Pathway: xylan degradation
C:Superfamily: Streptomyces endo-1,4-beta-xylanase A; Streptomyces endo-1,4-beta-xy
C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradati

Query Match 35.7%; Score 30; DB 2; Length 25;
Best Local Similarity 54.5%; Pred. No. 1.1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 4;

QY 4 AIHQVEYFVG 14
| : | | | |
DB 7 AAHQSGYFVG 17

RESULT 2

A60286
heat-stable serine proteinase (EC 3.4.21.-) - Thermomonospora fusca (strain YX) (fr
N:Alternate names: YX-proteinase
C:Species: Thermomonospora fusca
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Dec-1996
C:Accession: A60286
R:Kristjansson, M.M.; Kinsella, J.E.
Int. J. Pept. Protein Res. 36, 201-207, 1990
A:Title: Heat stable proteinase from Thermomonospora fusca. Characterization as a s
A:Reference number: A60286; MUID:91107200; PMID:2132918
A:Accession: A60286
A:Molecule type: protein
A:Residues: 1-25 <KRI>
C:Keywords: extracellular protein; hydrolase; serine proteinase

Query Match 35.7%; Score 30; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 1.1e+02; Indels 0; Gaps 0;
Matches 4; Conservative 2; Mismatches 0;

QY 11 YFQDFG 16

Db 10 YFYGNY 15
||||:

RESULT 3

PH1368

Ig heavy chain DJ region (clone C111-112) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: PH1368

R:Wasserman, R.; Galli, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.

J. Exp. Med. 176, 1577-1581, 1992

A:Title: Predominance of fetal type DJH joining in young children with B precursor lymph

A:Reference number: PH1302; MUID:93094761; PMID:1460419

A:Accession: PH1368

A:Molecule type: DNA

A:Residues: 1-18 <WAS>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 34.5%; Score 29; DB 2; Length 18;

Best Local Similarity 66.7%; Pred. NO. 1.2e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 9 VEYFG 14

:||||

Db 5 MEYFG 10

RESULT 4

PH1725

Ig heavy chain V region (clone GCC-5) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999

C:Accession: PH1725

R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.

J. Exp. Med. 178, 295-307, 1993

A:Title: Antigen-driven B cell differentiation in vivo.

A:Reference number: PH1675; MUID:93301607; PMID:8315385

A:Accession: PH1725

A:Molecule type: mRNA

A:Residues: 1-23 <MCH>

A:Experimental source: B cell

A>Note: The authors translated the codon ACA for residue 13 as Ala

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 33.3%; Score 28; DB 2; Length 23;

Best Local Similarity 33.3%; Pred. NO. 2.4e+02;

Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 8 QVEYFGDF 16

:||||

Db 12 ETRYFGSY 20

RESULT 5

PH1696

Ig heavy chain V region (clone NP-7-9) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999

C:Accession: PH1696

R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.

J. Exp. Med. 178, 295-307, 1993

A:Title: Antigen-driven B cell differentiation in vivo.

A:Reference number: PH1675; MUID:93301607; PMID:8315385

A:Accession: PH1696

A:Molecule type: mRNA

A:Residues: 1-24 <MCH>

A:Experimental source: B cell

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 33.3%; Score 28; DB 2; Length 24;
Best Local Similarity 57.1%; Pred. NO. 2.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 QVEYFG 14

:||||

Db 12 EVAYFG 18

RESULT 6

PH0771

T-cell receptor beta chain (PE5.1.1) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999

C:Accession: PH0771

R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.

J. Exp. Med. 174, 1371-1383, 1991

A:Title: T cell receptor genes in a series of class I major histocompatibility compl

A:Reference number: PH0746; MUID:92078846; PMID:1836010

A:Accession: PH0771

A:Molecule type: mRNA

A:Residues: 1-12 <CAS>

A:Cross-references: EMBL:X60865; NID:G53624; PIDN:CAA43255.1; PID:G53625

A:Experimental source: T lymphocyte

C:Keywords: T-cell receptor

Query Match 32.1%; Score 27; DB 2; Length 12;

Best Local Similarity 71.4%; Pred. NO. 1.8e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 QVEYFG 14

:||||

Db 6 QVEYFG 12

RESULT 7

A05305

hemoglobin beta-2 chain - Indian spiny-tailed lizard (fragment)

C:Species: Uromastix hardwickii (Indian spiny-tailed lizard)

C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 01-Aug-1997

C:Accession: A05305

R:Nagy, S.; Zaidi, Z.H.; von Bahr-Lindstrom, H.; Carlquist, M.; Jornvall, H.

FEBS Lett. 162, 290-295, 1983

A:Reference number: A91314; MUID:84029159; PMID:6628672

A:Accession: A05305

A:Molecule type: protein

A:Residues: 1-19 <NAO>

C:Superfamily: globin; globin homology

C:Keywords: erythrocyte; oxygen carrier

Query Match 32.1%; Score 27; DB 2; Length 19;

Best Local Similarity 80.0%; Pred. NO. 2.9e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 YFGDF 16

:||||

Db 1 FFGDF 5

RESULT 8

B53145

high conductance calcium-activated potassium channel, maxi-K channel - bovine (fragm

C:Species: Bos primigenius taurus (cattle)

C:Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 18-Jul-2001

C:Accession: B53145

R:Knaus, H.G.; Garcia-Calvo, M.; Kaczorowski, G.J.; Garcia, M.L.

J. Biol. Chem. 269, 3921-3924, 1994

A:Title: Subunit composition of the high conductance calcium-activated potassium ch

A:Reference number: A53145; MUID:94140798; PMID:7508434

A:Accession: B53145

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-19 <KNA>
A:Note: sequence extracted from NCBI backbone (NCBIP:144547)
C:Superfamily: fruit fly calcium-activated potassium channel slo

Query Match 32.1%; Score 27; DB 2; Length 19;
Best Local Similarity 71.4%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 QVEYFYG 14
: |||:
Db 4 QVEFYQG 10

RESULT 9
S46550
actin-related protein - chicken (fragments)
C:Species: Gallus gallus (Chicken)
C:Date: 12-Sep-1995 #sequence_revision 10-Nov-1995 #text_change 17-Mar-1999
C:Accession: S46550
R:Lees-Miller, J.P.; Helfman, D.M.; Schroer, T.A.
Nature 359, 244-246, 1992
A:Title: A vertebrate actin-related protein is a component of a multisubunit complex involved in the regulation of the actin cytoskeleton
A:Reference number: S29089; PMID:1528266
A:Accession: S46550
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-21 <LEE>

Query Match 32.1%; Score 27; DB 2; Length 21;
Best Local Similarity 44.4%; Pred. No. 3.2e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 ROVEYFYGD 15
: |||:
Db 13 RRAQYFLXD 21

RESULT 10
E53284
T-cell receptor beta 2 chain J region, Jbeta2.2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: E53284
R:Harindranath, N.; Alexander, C.B.; Mage, R.G.
Mol. Immunol. 28, 881-888, 1991
A:Title: Evolutionarily conserved organization and sequences of germline diversity and J region of the T-cell receptor beta 2 chain
A:Reference number: A53284; PMID:91342695; PMID:1678859
A:Accession: E53284
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-16 <HAR>
A:Cross-references: GB:S60737; NID:q233916; PIDN:AAB19521.1; PID:q233921
A:Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBIP:60743)
C:Keywords: T-cell receptor

Query Match 31.0%; Score 26; DB 2; Length 16;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 EYFQGD 15
: |||:
Db 4 QLYFGD 9

RESULT 11
PHI325
Ig heavy chain DJ region (clone C199-121) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PHI325
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma

A:Reference number: PHI302; PMID:93094761; PMID:1460419
A:Accession: PHI325
A:Molecule type: DNA
A:Residues: 1-22 <WAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 31.0%; Score 26; DB 2; Length 22;
Best Local Similarity 50.0%; Pred. No. 5e+02;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 VEYFYG 14
: |||:
Db 9 IHYYG 14

RESULT 12
PHI366
Ig heavy chain DJ region (clone C111-106) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PHI366
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma
A:Reference number: PHI302; PMID:93094761; PMID:1460419
A:Accession: PHI366
A:Molecule type: DNA
A:Residues: 1-15 <WAS>
C:Keywords: heterotetramer; immunoglobulin

Query Match 29.8%; Score 25; DB 2; Length 15;
Best Local Similarity 60.0%; Pred. No. 5e+02;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 YFGDF 16
: |||:
Db 6 YYGDY 10

RESULT 13
S57519
T cell receptor beta chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 10-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57519
R:Burrows, S.R.; Sillins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaet, V.P.
submitted to the EMBL Data Library, June 1995
A:Description: T cell receptor repertoire for a viral epitope in humans is diversified
A:Reference number: S57494
A:Accession: S57519
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-17 <BUR>
A:Cross-references: EMBL:249930; NID:g887494; PIDN:CAA90176.1; PID:g887495
C:Keywords: T-cell receptor

Query Match 29.8%; Score 25; DB 2; Length 17;
Best Local Similarity 55.6%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 IRQVEYFYG 14
: |||:
Db 9 ISSVEQYFG 17

RESULT 14
S57556
T cell receptor beta chain V-D-J region (clone PP7 and clone TF1) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57556
R:Burrows, S.R.; Sillins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaet, V.P.

submitted to the EMBL Data Library, June 1995
 A:Description: T cell receptor repertoire for a viral epitope in humans is diversified b
 A:Reference number: S57494
 A:Accession: S57556
 A:Molecule type: mRNA
 A:Residues: 1-17 <BUR>
 A:Cross-references: EMBL:Z49926; NID:9887498; PIDN:CAA90172.1; PID:9887499
 A:Experimental source: clone PP7
 A:Accession: S57557
 A:Molecule type: mRNA
 A:Residues: 1-17 <BUW>
 A:Cross-references: EMBL:Z49928; NID:9887502; PIDN:CAA90174.1; PID:9887503
 A:Experimental source: clone TFI
 C:Keywords: T-cell receptor

Query Match 29.8%; Score 25; DB 2; Length 17;
 Best Local Similarity 55.6%; Pred. No. 5.7e+02;
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 IRQVEYYEG 14
 | | | | |
 Db 9 ISSYEQYEG 17

RESULT 15

A61211
 anantin - Streptomyces coeruleus
 C:Species: Streptomyces coeruleus
 C:Date: 03-May-1994 #sequence_revision 05-Apr-1995 #text_change 07-May-1999
 C:Accession: A61211
 R:Wyss, D.F.; Lahm, H.W.; Manneberg, M.; Labhardt, A.M.
 J. Antibiot. 44, 172-180, 1991
 A:Title: Anantin -- a peptide antagonist of the atrial natriuretic factor (ANF). II. Det
 A:Reference number: A61211; MUID:91185186; PMID:1826288
 A:Accession: A61211
 A:Molecule type: protein
 A:Residues: 1-17 <WYS>
 A:Note: the isopeptide linked residue 8 is shown as Asn rather than Asp
 F:1-8/cross-link: isopeptide amino end (Gly-Asn) #status experimental

Query Match 29.8%; Score 25; DB 2; Length 17;
 Best Local Similarity 66.7%; Pred. No. 5.7e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 YFGDF 16
 : | | | |
 Db 12 HYGDF 17

Search completed: April 23, 2003, 13:48:56
 Job time : 8.35955 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:43:09 ; Search time 8.80999 Seconds
(without alignments)
145.542 Million cell updates/sec

Title: US-09-836-073-19
Perfect score: 84
Sequence: 1 QERAIQVEYFGDF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 61792

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap:*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	84	100.0	16	9	US-09-836-073-19
2	62	73.8	18	9	US-09-836-073-4
3	58	69.0	18	9	US-09-836-073-2
4	57	67.9	17	9	US-09-836-073-13
5	57	67.9	18	9	US-09-836-073-1
6	57	67.9	18	9	US-09-836-073-5
7	57	67.9	18	9	US-09-836-073-14
8	54	64.3	18	9	US-09-836-073-15
9	51	60.7	18	9	US-09-836-073-9
10	50	59.5	18	9	US-09-836-073-17
11	49	58.3	18	9	US-09-836-073-11
12	49	58.3	18	9	US-09-836-073-12
13	48	57.1	18	9	US-09-836-073-10
14	46.5	55.4	19	9	US-09-836-073-16
15	45	53.6	18	9	US-09-836-073-3
16	44.5	53.0	18	9	US-09-836-073-18
17	42	50.0	18	9	US-09-836-073-7
18	41	48.8	18	9	US-09-836-073-8
19	32	38.1	23	9	US-09-813-153-291

20	31	36.9	13	9	US-09-880-748-3091	Sequence 3091, Ap
21	31	36.9	23	9	US-10-012-542-421	Sequence 421, App
22	30	35.7	20	9	US-09-924-102-5	Sequence 5, Appl
23	29	34.5	18	9	US-10-084-813-139	Sequence 139, App
24	28	33.3	18	9	US-09-880-748-3060	Sequence 3060, Ap
25	28	33.3	18	9	US-10-084-813-140	Sequence 140, App
26	28	33.3	18	9	US-10-084-813-141	Sequence 141, App
27	28	33.3	18	10	US-09-840-009-27	Sequence 27, Appl
28	28	33.3	21	9	US-09-880-748-3010	Sequence 3010, Ap
29	27	32.1	15	10	US-09-073-009-51	Sequence 51, Appl
30	27	32.1	15	10	US-09-023-588-51	Sequence 51, Appl
31	27	32.1	15	10	US-09-793-306-51	Sequence 51, Appl
32	27	32.1	19	9	US-09-880-748-2987	Sequence 2987, Ap
33	27	32.1	21	10	US-09-764-017-2	Sequence 2, Appl
34	27	32.1	21	10	US-09-764-017-3	Sequence 3, Appl
35	27	32.1	25	10	US-09-782-980-150	Sequence 150, App
36	26	31.0	15	9	US-09-880-748-2886	Sequence 2886, Ap
37	26	31.0	15	9	US-09-880-748-2976	Sequence 2976, Ap
38	26	31.0	16	9	US-09-965-536A-56	Sequence 56, Appl
39	26	31.0	18	9	US-09-836-073-6	Sequence 6, Appl
40	26	31.0	19	9	US-09-880-748-2170	Sequence 2170, Ap
41	26	31.0	21	10	US-09-815-306-8	Sequence 8, Appl
42	26	31.0	25	10	US-09-864-761-43029	Sequence 43029, A
43	25.5	30.4	14	9	US-09-850-336-5	Sequence 2, Appl
44	25.5	30.4	18	9	US-09-850-336-5	Sequence 5, Appl
45	25.5	30.4	21	9	US-09-836-433-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-836-073-19
; Sequence 19, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila
US-09-836-073-19

Query Match 100.0%; Score 84; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QERAIQVEYFGDF 16
| | | | | | | | | | | | | | | |

Db 1 QERAIQVEYFGDF 16
| | | | | | | | | | | | | | | |

RESULT 2
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-4

Query Match 73.8%; Score 62; DB 9; Length 18;
 Best Local Similarity 68.8%; Pred. No. 0.00024;
 Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QERAIROVEYFGDF 16
 | | | | | | | | | | | | | | | | | |
 Db 3 QEAKICHOIEYFGDF 18

RESULT 3
 US-09-836-073-2
 ; Sequence 2, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 220002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-2

Query Match 69.0%; Score 58; DB 9; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.0012;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ERAIRQVEYFGDF 16
 | | | | | | | | | | | | | | | | | |
 Db 4 EAQICQIEYFGDF 18

RESULT 4
 US-09-836-073-13
 ; Sequence 13, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 220002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 17
 ; TYPE: PRT
 ; ORGANISM: Mouse
 US-09-836-073-13

Query Match 67.9%; Score 57; DB 9; Length 17;
 Best Local Similarity 66.7%; Pred. No. 0.0016;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ERAIRQVEYFGDF 16
 | | | | | | | | | | | | | | | | | |
 Db 3 EAKICHOIEYFGDF 17

RESULT 5
 US-09-836-073-1
 ; Sequence 1, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 220002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-1

Query Match 67.9%; Score 57; DB 9; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.0017;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ERAIRQVEYFGDF 16
 | | | | | | | | | | | | | | | | | |
 Db 4 EAKICHOIEYFGDF 18

RESULT 6
 US-09-836-073-5
 ; Sequence 5, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 220002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-5

Query Match 67.9%; Score 57; DB 9; Length 18;
 Best Local Similarity 62.5%; Pred. No. 0.0017;
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QERAIROVEYFGDF 16
 | | | | | | | | | | | | | | | | | |
 Db 3 QEQKCHOIEYFGDF 18

RESULT 7
 US-09-836-073-14

Thu Apr 24 08:54:00 2003

```
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9
```

```
Query Match 60.7%; Score 51; DB 9; Length 18;
Best Local Similarity 64.3%; Pred. No. 0.019;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 2 ERAIRQVEYYFGD 15
Db 4 EAKICHQIEYFGD 17
```

```
RESULT 10
US-09-836-073-17
; Sequence 17, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 18
; TYPE: PRT
; ORGANISM: C. elegans
US-09-836-073-17
```

```
Query Match 59.5%; Score 50; DB 9; Length 18;
Best Local Similarity 57.1%; Pred. No. 0.028;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 2 ERAIRQVEYYFGD 15
Db 4 DQRIKQLEYFGN 17
```

```
RESULT 11
US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11
```

```
Query Match 58.3%; Score 49; DB 9; Length 18;
Best Local Similarity 60.0%; Pred. No. 0.041;
```

```
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14
```

```
Query Match 67.9%; Score 57; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.0017;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 2 ERAIRQVEYYFGD 16
Db 4 EAKICHQIEYFGD 18
```

```
RESULT 8
US-09-836-073-15
; Sequence 15, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Xenopus
US-09-836-073-15
```

```
Query Match 64.3%; Score 54; DB 9; Length 18;
Best Local Similarity 75.0%; Pred. No. 0.0057;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 5 IIRQVEYYFGD 16
Db 7 ICEQIEYFGD 18
```

```
RESULT 9
US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
```

Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 ERAIRQVEYFGDF 16
| | | | |
Db 4 EAKICHQIEYFGDF 18

RESULT 12

US-09-836-073-12

; Sequence 12, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-12

Query Match

; Sequence 49; Score 49; DB 9; Length 18;

; Best Local Similarity 60.0%; Pred. No. 0.041;

Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 ERAIRQVEYFGDF 16
| | | | |
Db 4 EAKICHQIEYFGDF 18

RESULT 13

US-09-836-073-10

; Sequence 10, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-10

Query Match

; Sequence 48; Score 48; DB 9; Length 18;

; Best Local Similarity 60.0%; Pred. No. 0.061;

Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 ERAIRQVEYFGDF 16
| | | | |
Db 4 EAKICHQIEYFGDF 18

RESULT 14

US-09-836-073-16

; Sequence 16, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match

; Sequence 55.4%; Score 46.5; DB 9; Length 19;

; Best Local Similarity 62.5%; Pred. No. 0.12;

Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 2 ERAIRQV-EYFGDF 16
| | | | |
Db 4 EAKICHQIEYFGDF 19

RESULT 15

US-09-836-073-3

; Sequence 3, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-3

Query Match

; Sequence 53.6%; Score 45; DB 9; Length 18;

; Best Local Similarity 58.3%; Pred. No. 0.2;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 IIROVEYFGDF 16
| | | | |
Db 7 ICHOIYFGDF 18

Search completed: April 23, 2003, 13:52:12

Job time : 9.80899 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:38:28 ; Search time 7.55056 Seconds
(without alignments)
62.349 Million cell updates/sec

Title: US-09-836-073-19
Perfect score: 84
Sequence: 1 QERAIROVEYFGDF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 147762

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	67.9	18	US-09-316-630-3	Sequence 3, Appli
2	57	67.9	18	US-09-316-630-4	Sequence 4, Appli
3	35	41.7	11	US-08-618-696-7	Sequence 7, Appli
4	35	41.7	11	US-09-033-753-7	Sequence 7, Appli
5	32	38.1	11	US-08-618-696-1	Sequence 1, Appli
6	32	38.1	11	US-09-033-753-1	Sequence 1, Appli
7	31	36.9	11	US-08-618-696-6	Sequence 6, Appli
8	31	36.9	11	US-09-033-753-6	Sequence 6, Appli
9	30	35.7	10	US-08-618-696-11	Sequence 11, Appli
10	30	35.7	10	US-09-033-753-11	Sequence 11, Appli
11	30	35.7	11	US-08-618-696-20	Sequence 20, Appli
12	30	35.7	11	US-09-033-753-20	Sequence 20, Appli
13	30	35.7	25	US-08-453-289-4	Sequence 4, Appli
14	30	35.7	25	US-08-574-086-4	Sequence 4, Appli
15	29	34.5	14	US-09-298-924-55	Sequence 55, Appli
16	29	34.5	25	US-08-238-163-6	Sequence 6, Appli
17	28	33.3	10	US-08-618-696-2	Sequence 2, Appli
18	28	33.3	10	US-08-618-696-10	Sequence 10, Appli
19	28	33.3	10	US-09-033-753-10	Sequence 10, Appli
20	28	33.3	10	US-09-033-753-10	Sequence 10, Appli
21	28	33.3	11	US-08-618-696-16	Sequence 16, Appli
22	28	33.3	11	US-08-618-696-19	Sequence 19, Appli
23	28	33.3	11	US-09-033-753-16	Sequence 16, Appli
24	28	33.3	11	US-09-033-753-19	Sequence 19, Appli
25	27	32.1	8	US-08-444-818-530	Sequence 530, App
26	27	32.1	19	6 5196510-36	Patent No. 5196510
27	27	32.1	20	4 US-09-142-045A-2	Sequence 2, Appli

28	26	31.0	11	1	US-08-170-596-4	Sequence 4, Appli
29	26	31.0	14	2	US-08-726-464B-47	Sequence 47, Appli
30	26	31.0	16	1	US-08-218-025A-111	Sequence 111, App
31	26	31.0	18	6	5204445-9	Patent No. 5204445
32	26	31.0	20	1	US-08-127-351-21	Sequence 21, Appli
33	26	31.0	20	1	US-08-480-367B-21	Sequence 21, Appli
34	26	31.0	20	1	US-08-487-221A-21	Sequence 21, Appli
35	26	31.0	20	1	US-08-480-370-21	Sequence 21, Appli
36	26	31.0	20	6	5196510-29	Patent No. 5196510
37	26	31.0	20	6	5196510-30	Patent No. 5196510
38	26	31.0	21	1	US-07-965-674-15	Sequence 15, Appli
39	26	31.0	21	4	US-09-131-750-20	Sequence 20, Appli
40	26	31.0	21	5	PCT-US93-09523-15	Sequence 15, Appli
41	26	31.0	22	2	US-08-518-967-11	Sequence 11, Appli
42	26	31.0	25	1	US-08-486-421-8	Sequence 8, Appli
43	26	31.0	25	1	US-08-470-911-8	Sequence 8, Appli
44	26	31.0	25	2	US-08-486-809-8	Sequence 8, Appli
45	25.5	30.4	10	3	US-08-159-339A-665	Sequence 665, App

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match 67.9%; Score 57; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No: 0.00055;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIROVEYFGDF 16
| | | | | | | | | |
Db 4 EAKICHQIEYFGDF 18

RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

;; PRIOR FILING DATE: 1997-10-06
;; PRIOR APPLICATION NUMBER: 08/321,427
;; PRIOR FILING DATE: 1994-10-11
;; PRIOR APPLICATION NUMBER: 60/086,527
;; PRIOR FILING DATE: 1998-05-22
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
;; OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 67.9%; Score 57; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.00055;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERAIIQVEYFGDF 16
| | | | | | | | | |
Db 4 EAKICHOIEYFGDF 18

RESULT 3
US-08-618-696-7
; Sequence 7, Application US/08618696
; Patent No. 5861475
; GENERAL INFORMATION:
; APPLICANT: COOPER, JR., J. ALLEN D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/618,696
FILING DATE: 20-MAR-1996
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/995,269
FILING DATE: 12/21/92
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UOAB:002/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-618-696-7

Query Match 41.7%; Score 35; DB 2; Length 11;
Best Local Similarity 55.6%; Pred. No. 2.7;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 RQVEYFGD 15
| | | | |
Db 1 REASYFFGD 9

RESULT 4
US-09-033-753-7
; Sequence 7, Application US/09033753
; Patent No. 6017883
; GENERAL INFORMATION:
; APPLICANT: COOPER, JR., J. ALLEN D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,753
FILING DATE:
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/618,696
FILING DATE: 20-MAR-1996
APPLICATION NUMBER: 07/995,269
FILING DATE: 12/21/92
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UOAB:002/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-033-753-7

Query Match 41.7%; Score 35; DB 3; Length 11;
Best Local Similarity 55.6%; Pred. No. 2.7;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 RQVEYFGD 15
| | | | |
Db 1 REASYFFGD 9

RESULT 5
US-08-618-696-1
; Sequence 1, Application US/08618696
; Patent No. 5861475
; GENERAL INFORMATION:
; APPLICANT: COOPER, JR., J. ALLEN D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON

STATE: TEXAS
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/618,696
FILING DATE: 20-MAR-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/995,269
FILING DATE: 12/21/92
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UOAB:002/PAR
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-618-696-1

Query Match 38.1%; Score 32; DB 2; Length 11;
Best Local Similarity 55.6%; Pred. No. 9.4;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 RQVEYFEGD 15
I: |:|:|
Db 1 REGSYFFGD 9

RESULT 6
US-09-033-753-1
Sequence 1, Application US/09033753
Patent No. 6017883
GENERAL INFORMATION:
APPLICANT: COOPER, JR., J. ALLEN D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,753
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/618,696
FILING DATE: 20-MAR-1996
APPLICATION NUMBER: 07/995,269
FILING DATE: 12/21/92
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165

REFERENCE/DOCKET NUMBER: UOAB:002/PAR
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-033-753-1

Query Match 38.1%; Score 32; DB 3; Length 11;
Best Local Similarity 55.6%; Pred. No. 9.4;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 RQVEYFEGD 15
I: |:|:|
Db 1 REGSYFFGD 9

RESULT 7
US-08-618-696-6
Sequence 6, Application US/08618696
Patent No. 5861475
GENERAL INFORMATION:
APPLICANT: COOPER, JR., J. ALLEN D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/618,696
FILING DATE: 20-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/995,269
FILING DATE: 12/21/92
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UOAB:002/PAR
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-618-696-6

Query Match 36.9%; Score 31; DB 2; Length 11;
Best Local Similarity 55.6%; Pred. No. 14;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 RQVEYFEGD 15
I: |:|:|
Db 1 REGSYFFGD 9


```

, , COMPUTER READABLE FORM:
, , MEDIUM TYPE: FLOPPY DISK
, , COMPUTER: IBM PC COMPATIBLE
, , OPERATING SYSTEM: PC-DOS/MS-DOS
, , SOFTWARE: WORDPERFECT 5.1
, , CURRENT APPLICATION DATA:
, , APPLICATION NUMBER: US/08/618,696
, , FILING DATE: 20-MAR-1996
, , CLASSIFICATION: 514
, , PRIOR APPLICATION NUMBER:
, , FILING DATE:
, , CLASSIFICATION:
, , APPLICATION NUMBER: 07/995,269
, , FILING DATE: 12/21/92
, , ATTORNEY/AGENT INFORMATION:
, , NAME: PARKER, DAVID L.
, , REGISTRATION NUMBER: 32,165
, , REFERENCE/DOCKET NUMBER: UOAB:002/PAR
, , TELECOMMUNICATION INFORMATION:
, , TELEPHONE: 512-320-7200
, , TELEFAX: 512-474-7577
, , TELEX: NOT APPLICABLE
, , INFORMATION FOR SEQ ID NO: 11:
, , SEQUENCE CHARACTERISTICS:
, , LENGTH: 10 amino acid residues
, , TYPE: amino acid
, , STRANDEDNESS: single
, , TOPOLOGY: linear
, ,
, , US-08-618-696-11
, ,
, , Query Match 35.7%; Score 30;
, , Best Local Similarity 50.0%; Pred. No. 1;
, , Matches 4; Conservative 2; Mismatchch
, ,
, , QY 8 QVEYVFGD 15
, , : I:III
, , Db 1 EASYFGD 8
, ,
, , RESULT 10
, , US-09-033-753-11
, , Sequence 11, Application US/09033753
, , Patent No. 6017883
, , GENERAL INFORMATION:
, , APPLICANT: COOPER, Jr., J. ALLEN D.
, , TITLE OF INVENTION: COMPOSITIONS AND MET
, , TITLE OF INVENTION: INHIBITION OF PHAGO
, , NUMBER OF SEQUENCES: 21
, , CORRESPONDENCE ADDRESS:
, , ADDRESSEE: ARNOLD, WHITE & DURKEE
, , STREET: P.O. BOX 4433
, , CITY: HOUSTON
, , STATE: TEXAS
, , COUNTRY: USA
, , ZIP: 77210
, ,
, , COMPUTER READABLE FORM:
, , MEDIUM TYPE: FLOPPY DISK
, , COMPUTER: IBM PC COMPATIBLE
, , OPERATING SYSTEM: PC-DOS/MS-DOS
, , SOFTWARE: WORDPERFECT 5.1
, , CURRENT APPLICATION DATA:
, , APPLICATION NUMBER: US/09/033,753
, , FILING DATE:
, , CLASSIFICATION: 530
, , PRIOR APPLICATION DATA:
, , APPLICATION NUMBER: 08/618,696
, , FILING DATE: 20-MAR-1996
, , APPLICATION NUMBER: 07/995,269
, , FILING DATE: 12/21/92
, , ATTORNEY/AGENT INFORMATION:
, , NAME: PARKER, DAVID L.
, , REGISTRATION NUMBER: 32,165
, , REFERENCE/DOCKET NUMBER: UOAB:002/PAR
, , TELECOMMUNICATION INFORMATION:
, , TELEPHONE: 512-320-7200

```

TELEFAX: 512-474-7577
 TELEX: NOT APPLICABLE
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acid residues
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-033-753-11

Query Match 35.7%; Score 30; DB 3; Length 10;
 Best Local Similarity 50.0%; Pred. No. 19;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 QVEYFFGD 15
 : |:|
 Db 1 EASYFFGD 8

RESULT 11
 US-08-618-696-20
 ; Sequence 20, Application US/08618696
 ; Patent No. 5861475
 ; GENERAL INFORMATION:
 ; APPLICANT: COOPER, JR., J. ALLEN D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
 ; INHIBITION OF PHAGOCYTES
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ARNOLD, WHITE & DURKEE
 ; STREET: P.O. BOX 4433
 ; CITY: HOUSTON
 ; STATE: TEXAS
 ; COUNTRY: USA
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WORDPERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/618,696
 ; FILING DATE: 20-MAR-1996
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/995,269
 ; FILING DATE: 12/21/92
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PARKER, DAVID L.
 ; REGISTRATION NUMBER: 32,165
 ; REFERENCE/DOCKET NUMBER: UOAB:002/PAR
 ; TELEPHONE: 512-320-7200
 ; TELEFAX: 512-474-7577
 ; TELEX: NOT APPLICABLE
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 11 amino acid residues
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-618-696-20

Query Match 35.7%; Score 30; DB 2; Length 11;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 QVEYFFGD 15
 : |:|
 Db 2 EASYFFGD 9

RESULT 12

US-09-033-753-20
 ; Sequence 20, Application US/09033753
 ; Patent No. 6017883
 ; GENERAL INFORMATION:
 ; APPLICANT: COOPER, JR., J. ALLEN D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
 ; INHIBITION OF PHAGOCYTES
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: ARNOLD, WHITE & DURKEE
 ; STREET: P.O. BOX 4433
 ; CITY: HOUSTON
 ; STATE: TEXAS
 ; COUNTRY: USA
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC COMPATIBLE
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WORDPERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/033,753
 ; FILING DATE:
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/618,696
 ; FILING DATE: 20-MAR-1996
 ; APPLICATION NUMBER: 07/995,269
 ; FILING DATE: 12/21/92
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PARKER, DAVID L.
 ; REGISTRATION NUMBER: 32,165
 ; REFERENCE/DOCKET NUMBER: UOAB:002/PAR
 ; TELEPHONE: 512-320-7200
 ; TELEFAX: 512-474-7577
 ; TELEX: NOT APPLICABLE
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 11 amino acid residues
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-09-033-753-20

Query Match 35.7%; Score 30; DB 3; Length 11;
 Best Local Similarity 50.0%; Pred. No. 21;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 QVEYFFGD 15
 : |:|
 Db 2 EASYFFGD 9

RESULT 13
 US-08-453-289-4
 ; Sequence 4, Application US/08453289
 ; Patent No. 5498534
 ; GENERAL INFORMATION:
 ; APPLICANT: Jeffries, Thomas W
 ; APPLICANT: Grabski, Anthony C
 ; APPLICANT: Patel, Rajesh N.
 ; APPLICANT: Elegir, Graziano
 ; APPLICANT: Szakacs, George
 ; TITLE OF INVENTION: Method of Removing Color from Kraft Wood
 ; PULPS
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Janet I. Stockhausen
 ; STREET: One Gifford Pinchot Drive
 ; CITY: Madison
 ; STATE: Wisconsin
 ; COUNTRY: USA

;; ZIP: 53705
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/453,289
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/257,965
;; FILING DATE:
;; APPLICATION NUMBER: US 07/857,060
;; FILING DATE: 25-MAY-1992
;; NAME: Stockhausen, Janet I
;; REGISTRATION NUMBER: 34,256
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 608-231-9504
;; TELEFAX: 608-231-9508
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 25 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;; ORIGINAL SOURCE:
;; ORGANISM: Streptomyces roseiscleroticus
;; STRAIN: NRRLB-11019
;; US-08-453-289-4

Query Match 35.7%; Score 30; DB 1; Length 25;
Best Local Similarity 54.5%; Pred. No. 55;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 AIIROVEYFYG 14
| :| ||||
Db 7 AAQOSGYFYG 17

RESULT 14
US-08-574-086-4
; Sequence 4, Application US/08574086
; Patent No. 5834301
; GENERAL INFORMATION:
; APPLICANT: Jeffries, Thomas W
; APPLICANT: Grabski, Anthony C
; APPLICANT: Patel, Rajesh N
; APPLICANT: Elegir, Graziano
; APPLICANT: Szakacs, George
; TITLE OF INVENTION: Method of Removing Color from Kraft Wood
; TITLE OF INVENTION: Pulp
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janet I. Stockhausen
; STREET: One Gifford Pinchot Drive
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53705
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,086
; FILING DATE: 18-DEC-1995
; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/453,289
;; FILING DATE:
;; APPLICATION NUMBER: US/08/257,965
;; FILING DATE:
;; APPLICATION NUMBER: US 07/857,060
;; FILING DATE: 25-MAY-1992
;; NAME: Stockhausen, Janet I
;; REGISTRATION NUMBER: 34,256
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 608-231-9504
;; TELEFAX: 608-231-9508
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 25 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: N-terminal
;; ORIGINAL SOURCE:
;; ORGANISM: Streptomyces roseiscleroticus
;; STRAIN: NRRLB-11019
;; US-08-574-086-4

Query Match 35.7%; Score 30; DB 2; Length 25;
Best Local Similarity 54.5%; Pred. No. 55;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 AIIROVEYFYG 14
| :| ||||
Db 7 AAQOSGYFYG 17

RESULT 15
US-09-298-924-55
; Sequence 55, Application US/09298924
; Patent No. 6391595
; GENERAL INFORMATION:

APPLICANT: KATO, Masaru
MIURA, Yutaka
KETTOU, Masako
IWAMATSU, Akihiro
KOBAYASHI, Kazuo
KOMEDA, Toshihiro

TITLE OF INVENTION: NOVEL TRANSFERASE AND AMYLASE, PROCESS
FOR PRODUCING THE ENZYMES, USE THEREOF, AND GENE CODING
FOR THE SAME

NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA

ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/298,924
FILING DATE: 26-Apr-1999
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/750,569
FILING DATE: <Unknown>
APPLICATION NUMBER: JP 7-120673
FILING DATE: 21-APR-1995
APPLICATION NUMBER: JP 6-311185

```

; FILING DATE: 21-NOV-1994
; APPLICATION NUMBER: JP 6-286917
; FILING DATE: 21-NOV-1994
; APPLICATION NUMBER: JP 6-290394
; FILING DATE: 31-OCT-1994
; APPLICATION NUMBER: JP 6-194223
; FILING DATE: 18-AUG-1994
; APPLICATION NUMBER: JP 6-133354
; FILING DATE: 16-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 49441/110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; LENGTH: 14 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-298-924-55
Query Match      34.5%  Score 29; DB 4; Length 14;
Best Local Similarity 57.1%  Pred. No. 43;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 10 EYFGDF 16
   :||: ||
Db 1 DYYQDF 7

```

Search completed: April 23, 2003, 13:50:29
Job time : 8.55036 secs

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:28:18 ; Search time 24.1798 Seconds
(without alignments)
88.173 Million cell updates/sec

Title: US-09-836-073-19

Perfect score: 84
Sequence: 1 QERAIIRQVEYFGDF 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 320064

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	67.9	18	AA52200	Human la autoantigen
2	54	64.3	21	AA43394	La/SSB epitope 17
3	35	41.7	11	AA56292	Synthetic derivat
4	32	38.1	11	AA56288	Native human neutr
5	32	38.1	11	AA56288	Chemotactic peptid
6	32	38.1	23	AA25901	Human secreted pro
7	31	36.9	8	AA43395	La/SSB epitope 24
8	31	36.9	10	AA22202	Murine MC-1 antibo
9	31	36.9	11	AA56291	Synthetic derivati
10	31	36.9	13	ABP47080	Human Blys binding

11	31	36.9	21	AA586493	Human gene 61-enco
12	31	36.9	24	AA19735	SEQ ID NO 453 from
13	30	35.7	10	AA56297	Synthetic modified
14	30	35.7	11	AA56307	Modified Influenza
15	30	35.7	20	AAE19842	Human hrp derived
16	30	35.7	25	AAE21213	Streptomyces rosei
17	30	35.7	25	AAW80367	N-terminal sequenc
18	29	34.5	14	AA90865	S. acidocaldarius
19	29	34.5	18	AA20426	Anti-FIX/FIXA anti
20	29	34.5	18	AA89021	HIV gp120 protein
21	29	34.5	23	AAU0924	VH ligand-binding
22	28	33.3	9	ABP47453	N. meningitidis LO
23	28	33.3	10	AA56293	Native human neutr
24	28	33.3	10	AA56296	Synthetic modified
25	28	33.3	11	AA56306	Modified Influenza
26	28	33.3	11	AA56303	Influenza A nucleo
27	28	33.3	11	AAW5255	Anti-progesterone
28	28	33.3	11	ABP47593	N. meningitidis LO
29	28	33.3	14	AAW6759	Human peptide #34
30	28	33.3	14	AAW7227	Human peptide #502
31	28	33.3	18	AA89022	HIV gp120 protein
32	28	33.3	18	AA89023	HIV gp120 protein
33	28	33.3	18	ABP47049	Human Blys binding
34	28	33.3	18	AAE20879	Ell-novispirin pep
35	28	33.3	20	AAU90008	Insulin/insulin-li
36	28	33.3	21	AA92191	Signal transductio
37	28	33.3	21	ABG66486	IgE Fc epsilon RI b
38	28	33.3	21	ABP46999	Human Blys binding
39	28	33.3	23	AA60984	Sequence of mumps
40	28	33.3	24	AA94731	Region B of protei
41	28	33.3	24	AA89801	Protein kinase der
42	27	32.1	8	AA36046	Hepatitis C virus
43	27	32.1	10	AAU02729	CDR region of anti
44	27	32.1	10	AAU83354	Antiviral composit
45	27	32.1	10	AAU83360	Antiviral composit

ALIGNMENTS

RESULT 1
AA52200
ID AA52200 standard; peptide; 18 AA.
XX AC AA52200;
DT 14-MAR-2000 (first entry)
XX DE Human la autoantigen peptide (LAP).

XX DE La autoantigen; LAP; internal ribosome entry site; IRES; translation;
KW viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
KW coronavirus; hepatitis virus; rhinovirus; adenovirus; coxsackie virus;
KW parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
KW vesicular stomatitis virus.

XX OS Homo sapiens.
XX PN WO9961613-A2.
XX PD 02-DEC-1999.
XX PF 21-MAY-1999; 99WO-US11281.
XX PR 22-MAY-1998; 98US-0086527.
XX PA (REGC) UNIV CALIFORNIA.
XX PI Das S, Dasgupta A;
XX DR WPI; 2000-062712/05.

PT New yeast inhibitory peptide useful for inhibiting viral protein
 PT translation and replication
 XX
 PS . Claim 5; Page 57; 81pp; English.
 XX
 CC This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (IRNA see A4245200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rhinovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.
 XX
 SQ Sequence 18 AA;

Query Match 67.98; Score 57; DB 21; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.0028;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 ERAIIROVEYFGDF 16
 Db 4 EAKICHQIEYFGDF 18

RESULT 2
 AAR43394
 ID AAR43394 standard; peptide; 21 AA.
 XX
 AC AAR43394;
 DT 12-MAY-1994 (first entry)
 DE La/SSB epitope 17.
 XX
 KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9321223-A.
 XX
 PD 28-OCT-1993.
 XX
 PF 13-APR-1993; 93WO-US03484.
 XX
 PR 13-APR-1992; 92US-0867819.
 XX
 PA (OKLA) UNIV OKLAHOMA STATE.
 XX
 PI Harley JB;
 XX
 DR WPI; 1993-351658/44.
 XX
 XX New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus
 XX
 PS Claim 1; Page 30; 43pp; English.
 XX
 CC The sequences given in AAR43391-562 are linear epitopes which are

CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polypyridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.
 XX
 SQ Sequence 21 AA;

Query Match 64.38; Score 54; DB 14; Length 21;
 Best Local Similarity 75.0%; Pred. No. 0.011;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 5 IIRQVEYFGDF 16
 Db 1 ICHQIEYFGDF 12

RESULT 3
 AAR56292
 ID AAR56292 standard; Peptide; 11 AA.
 XX
 AC AAR56292;
 DT 04-MAR-1995 (first entry)
 DE Synthetic derivative of human neutrophil inhibitor peptide (NIP).
 DE Neutrophil inhibitor peptide; NIP; phagocyte inhibitor;
 DE degranulation inhibitor; pulmonary inflammation.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Active-site 3..5
 FT Active-site /label= modified phosphorylation site
 FT 3..7 /label= modified core sequence
 XX
 PN WO9414463-A.
 XX
 PD 07-JUL-1994.
 XX
 PF 20-DEC-1993; 93WO-US12474.
 XX
 PR 21-DEC-1992; 92US-0995269.
 XX
 PA (UABR-) UAB RES FOUND.
 XX
 PI Cooper JAD;
 XX
 DR WPI; 1994-234345/28.
 XX
 XX Novel neutrophil inhibitory peptides - inhibit phagocyte
 PT function, used to reduce pulmonary inflammation
 PT
 XX Claim 15; Page 15; 97pp; English.
 PS
 XX Bronchoalveolar lavage was performed on lungs obtd. from heart
 CC transplant donors. At the time of death there was no evidence of

CC lung infection nor previous lung disease. Inhibitor peptides were
 CC purified from one litre of lavage fluid. Sequence analysis revealed
 CC two peptides, called native NIP (AAR56288) and NIP-Arg (AAR56293). The
 CC sequences exhibited a striking homology to a stretch of AAs in the
 CC sequence of a nucleoprotein prod. by certain Influenza A viruses
 CC (AAR56303). Partial homology to sequences within myf-6, ros and neu
 CC oncogene-related proteins was also observed. NIP and NIP-Arg
 CC contain a potential phosphorylation site comprising G-S-Y; either
 CC or both of which S and Y residues have the potential to become
 CC phosphorylated. Such phosphorylation may result in the reduction
 CC of the inhibitory action of the peptides. A truncated synthetic
 CC peptide (AAR56298) with the core sequence of NIP, namely G-S-Y-F-F.
 CC residues S, Y and also G, which form both the target and additional
 CC recognitions elements of the phosphorylation site. Such
 CC peptide include the second generation sequences
 CC in AAR56289-92; AAR56294-97; AAR56299-302; and AAR56304-07. AAR56308 is
 CC a peptide with the same proportion of AAs as NIP but in a random
 CC sequence. It was produced for use as a control in cellular
 CC function studies.

XX SQ Sequence 11 AA;

Query Match 41.7%; Score 35; DB 15; Length 11;
 Best Local Similarity 55.6%; Pred. No. 12;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 RQVEYFFGD 15
 |:|:|:|
 Db 1 REASYFFGD 9

RESULT 4
 AAR56288
 ID AAR56288 standard; Peptide; 11 AA.
 AC AAR56288;
 XX 04-MAR-1995 (first entry)
 DT Native human neutrophil inhibitor peptide (NIP).
 DE Neutrophil inhibitor peptide; NIP; phagocyte inhibitor;
 KW degranulation inhibitor; pulmonary inflammation.
 XX Homo sapiens.
 OS Key Location/Qualifiers
 FH Active-site 3..5
 FT /label= potential phosphorylation site
 FT Active-site 3..7
 FT /label= core sequence
 XX WO9414463-A.
 XX 07-JUL-1994.
 XX 20-DEC-1993; 93WO-US12474.
 XX 21-DEC-1992; 92US-0995269.
 XX (UABR-) UAB RES FOUND.
 XX Cooper JAD;
 XX WPI; 1994-234345/28.
 XX Novel neutrophil inhibitory peptides - inhibit phagocyte
 PT function, used to reduce pulmonary inflammation
 XX

PS Claim 9; Page 15; 97pp; English.
 XX Bronchoalveolar lavage was performed on lungs obtd. from heart
 CC transplant donors. At the time of death there was no evidence of
 CC lung infection nor previous lung disease. Inhibitor peptides were
 CC purified from one litre of lavage fluid. Sequence analysis revealed
 CC two peptides - native NIP (AAR56288) and native NIP-Arg (AAR56293). The
 CC sequences exhibited a striking homology to a stretch of AAs in the
 CC sequence of a nucleoprotein prod. by certain Influenza A viruses
 CC (AAR56303). Partial homology to sequences within myf-6, ros and neu
 CC oncogene-related proteins was also observed. NIP and NIP-Arg
 CC contain a potential phosphorylation site comprising G-S-Y; either
 CC or both of which S and Y residues have the potential to become
 CC phosphorylated. Such phosphorylation may result in the reduction
 CC of the inhibitory action of the peptides. A truncated synthetic
 CC peptide (AAR56298) with the core sequence of NIP, namely G-S-Y-F-F.
 CC was found to inhibit PMN chemotaxis with approx. the same potency
 CC as NIP. The invention pertains to peptides which inhibit PMN
 CC activation and include within their sequence the sequence AAR56298,
 CC or AAR56293, AAR56288; but which contain modifications in those
 CC residues S, Y and also G, which form both the target and additional
 CC recognitions elements of the phosphorylation site. Such
 CC peptide include the second generation sequences
 CC in AAR56289-92; AAR56294-97; AAR56299-302; and AAR56304-07. AAR56308 is
 CC a peptide with the same proportion of AAs as NIP but in a random
 CC sequence. It was produced for use as a control in cellular
 CC function studies.

XX SQ Sequence 11 AA;

Query Match 38.1%; Score 32; DB 15; Length 11;
 Best Local Similarity 55.6%; Pred. No. 41;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 RQVEYFFGD 15
 |:|:|:|
 Db 1 REGSYFFGD 9

RESULT 5
 AAB91936
 ID AAB91936 standard; Peptide; 11 AA.
 AC AAB91936;
 XX 22-JUN-2001 (first entry)
 DT Chemotactic peptide SEQ ID NO:1112.
 DE Protection; endogenous therapeutic peptide; peptidase; conjugation;
 KW blood component; modification; succinimide; maleimido group; amino;
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
 XX Homo sapiens.
 OS Synthetic.
 XX WO2000069900-A2.
 XX 23-NOV-2000.
 XX 17-MAY-2000; 2000WO-US13576.
 XX 17-MAY-1999; 99US-0134406.
 XX 10-SEP-1999; 99US-0153406.
 XX 15-OCT-1999; 99US-0159783.
 XX (CONJ-) CONJUCHEM INC.
 XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;
 XX WPI; 2001-112059/12.
 XX Modifying and attaching therapeutic peptides to albumin prevents
 PT

PT peptidase degradation, useful for increasing length of in vivo activity
 PT
 XX
 PS Disclosure; Page 559; 733pp; English.
 XX
 CC The present invention describes a modified therapeutic peptide (I)
 CC comprising a therapeutically active amino acid region (III) and a
 CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
 CC a less therapeutically active amino acid region (IV), which covalently
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity
 CC in vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or
 CC reduces the action of peptidases to increase length of activity (half
 CC life) and specificity as bonding to large molecules decreases
 CC intracellular uptake and interference with physiological processes.
 CC AAB90829 to AAB92441 represent peptides which can be used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 11 AA;
 Query Match 38.1%; Score 32; DB 22; Length 11;
 Best Local Similarity 55.6%; Pred. No. 41;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 7 ROVEYFCD 15
 I: | | | | |
 Db 1 REGSYFFGD 9
 RESULT 6
 AAY25901
 ID AAY25901 standard; Protein; 23 AA.
 XX
 AC AAY25901;
 XX
 DT 04-OCT-1999 (first entry)
 XX
 DE Human secreted protein fragment encoded from gene 64.
 XX
 KW Secreted protein; human; treatment; diagnosis; therapy; cancer; tumour;
 KW neurodegenerative disorder; developmental abnormality; blood disorder;
 KW fetal deficiency; blood disorder; leukemia; immune system; inflammation;
 KW autoimmune disease; hepatic disease; renal disease; allergy; restenosis;
 KW ischaemic shock; Alzheimer's disease; cognitive disorder; schizophrenia;
 KW cardiovascular disorder; wound healing; stroke; arthritis; obesity;
 KW asthma; sepsis; acne; psoriasis; transplant rejection; infection; AIDS;
 KW metabolic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO9938881-A1.
 XX
 PD 05-AUG-1999.
 XX
 PF 27-JAN-1999; 99WO-US01621.
 XX
 PR 30-JAN-1998; 98US-00731170.
 PR 30-JAN-1998; 98US-00731159.
 PR 30-JAN-1998; 98US-00731160.
 PR 30-JAN-1998; 98US-00731161.
 PR 30-JAN-1998; 98US-00731162.
 PR 30-JAN-1998; 98US-00731164.
 PR 30-JAN-1998; 98US-00731165.
 PR 30-JAN-1998; 98US-00731167.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Carter KC, Endress GA, Feng P, Ferrie AM, Florence C;
 PI

PI Florence KA, Janat F, Ni J, Rosen CA, Ruben SM;
 PI Soppet DR, Young P, Yu G;
 XX
 DR WPI; 1999-469315/39.
 DR N-PSDB; AA200473.
 XX
 XX New isolated human genes and the secreted polypeptides they encode
 PT useful in, e.g. treatment of Alzheimer's
 XX
 PS Disclosure; Page 382; 393pp; English.
 XX
 CC This invention describes novel human genes (see AA200410-200477) and the
 CC secreted proteins (see AAY25711-Y25778) and fragments (see
 CC AAY25779-Y25907) they encode. The polynucleotides and their corresponding
 CC secreted polypeptides are useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. Also pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 67
 CC polynucleotides of the invention, based on which tissues they are most
 CC highly expressed in, and include developing products for the diagnosis
 CC or treatment of cancer, tumours, neurodegenerative disorders,
 CC developmental abnormalities and fetal deficiencies, blood disorders,
 CC leukemias, diseases of the immune system, autoimmune diseases, hepatic
 CC and renal disease, inflammation, allergies, ischaemic shock, Alzheimer's
 CC and cognitive disorders, schizophrenia, restenosis, cardiovascular
 CC disorders, wound healing, stroke, arthritis, obesity, asthma, sepsis,
 CC acne, psoriasis, transplant rejection, metabolic disorders, infections
 CC and AIDS. The polypeptides are also useful for identifying their binding
 CC partners.
 XX
 SQ Sequence 23 AA;
 Query Match 38.1%; Score 32; DB 20; Length 23;
 Best Local Similarity 63.6%; Pred. No. 97;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 QERAIIRQVEY 11
 I: | | | | |
 Db 11 QERAKIHQTEH 21
 RESULT 7
 AAR43395
 ID AAR43395 standard; peptide; 8 AA.
 XX
 AC AAR43395;
 XX
 DT 12-MAY-1994 (first entry)
 XX
 DE La/SSB epitope 24.
 XX
 KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9321223-A.
 XX
 PD 28-OCT-1993.
 XX
 PF 13-APR-1993; 93WO-US03484.
 XX
 PR 13-APR-1992; 92US-0867819.
 XX
 PA (OKLA) UNIV OKLAHOMA STATE.
 XX
 PI Harley JB;
 XX
 DR WPI; 1993-351658/44.
 XX

PT New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus.
 XX
 PS Claim 1; Page 30; 43pp; English.
 XX
 CC The sequences given in AAR43391-562 are linear epitopes which are
 CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a pyrimidine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.

XX Sequence 8 AA;

Query Match 36.9%; Score 31; DB 14; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 YFGDF 16
 |||||
 Db 1 YFGDF 5

RESULT 8

AAE22202
 ID AAE22202 standard; peptide; 10 AA.

XX

AC AAE22202;

XX

DT 25-JUL-2002 (first entry)

XX

DE Murine MC-1 antibody heavy chain variable domain (VH) CDR3 peptide.

XX Chemokine construct; human immunodeficiency virus 1; allergic disease;
 KW skin disease; immunological disorder; autoimmune disease; psoriasis;
 KW multiple sclerosis; rheumatoid arthritis; inflammatory bowel disease;
 KW diabetes; skin inflammation; atopic dermatitis; inflammatory disease;
 KW inflammatory renal disease; HIV-1; transplant rejection; murine; CDR3;
 KW heavy chain variable domain; VH; complementarity determining region 3;
 KW MC-1; antibody.

XX

OS Mus sp.

XX

FN WO200220615-A2.

XX

PD 14-MAR-2002.

XX

PF 10-SEP-2001; 2001WO-EPI0433.

XX

PR 08-SEP-2000; 2000EP-0119694.

XX

PR 05-SEP-2001; 2001US-0948004.

XX

PA (MICR-) MICROMET AG.

XX

XX Mack M, Schloendorff D, Spring M;

XX

DR WPI; 2002-362240/39.

XX Use of an antibody and/or chemokine construct that binds to a chemokine
 PT receptor, for eliminating cells latently infected with primate
 PT immunodeficiency virus, or treating, preventing and alleviating immune
 PT disorders
 XX
 PS Disclosure; Page 117; 117pp; English.
 XX
 CC The invention relates to the use of an antibody and/or chemokine
 CC construct that binds a chemokine receptor for preparing a pharmaceutical
 CC composition for eliminating cells latently infected with a primate
 CC immunodeficiency virus (e.g. human immunodeficiency virus (HIV-1)). They
 CC are used in gene therapy and as vaccines. The antibody and/or chemokine
 CC construct is also used for preparing a pharmaceutical composition for
 CC treating, preventing and/or alleviating immunological disorders
 CC including autoimmune diseases (e.g. multiple sclerosis, type I diabetes
 CC and rheumatoid arthritis), allergic diseases, skin diseases (e.g. skin
 CC inflammation, atopic dermatitis and psoriasis), inflammatory diseases
 CC such as inflammatory joint diseases (chronic arthritis), inflammatory
 CC renal diseases and inflammatory bowel diseases and graft versus host
 CC and transplant rejections. The present sequence is murine MC-1 antibody
 CC heavy chain variable domain (VH) complementarity determining region 3
 CC (CDR3) peptide.

XX Sequence 10 AA;

Query Match 36.9%; Score 31; DB 23; Length 10;
 Best Local Similarity 71.4%; Pred. No. 56;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 10 EYFGDF 16
 |||||
 Db 2 EYVGIF 8

RESULT 9

AAE56291

ID AAR56291 standard; Peptide; 11 AA.

XX

AC AAR56291;

XX

DT 04-MAR-1995 (first entry)

XX

DE Synthetic derivative of human neutrophil inhibitor peptide (NIP).

XX

KW Neutrophil inhibitor peptide; NIP; phagocyte inhibitor;
 KW degranulation inhibitor; pulmonary inflammation.

XX

OS Synthetic.

XX

FN Key Location/Qualifiers

FT Active-site 3.5

FT Active-site /label- modified phosphorylation site

FT Active-site 3.7

FT Active-site /label- modified core sequence

XX

PN WO9414463-A.

XX

PD 07-JUL-1994.

XX

PF 20-DEC-1993; 93WO-US12474.

XX

PR 21-DEC-1992; 92US-0995269.

XX

PA (UABR-) UAB RES FOUND.

XX

PI Cooper JAD;

XX

DR WPI; 1994-234345/28.

XX

PT Novel neutrophil inhibitory peptides - inhibit phagocyte
 PT function, used to reduce pulmonary inflammation
 XX

PI Claim 14; Page 15; 97pp; English.

XX Bronchoalveolar lavage was performed on lungs obtd. from heart
CC transplant donors. At the time of death there was no evidence of
CC lung infection nor previous lung disease. Inhibitor peptides were
CC purified from one litre of lavage fluid. Sequence analysis revealed
CC two peptides - native NIP (AAR56288) and NIP-Arg (AAR56293). The
CC sequences exhibited a striking homology to a stretch of AAs in the
CC sequence of a nucleoprotein prod. by certain Influenza A viruses
CC (AAR56303). Partial homology to sequences within myf-6, ros and neu
CC oncogene-related proteins was also observed. NIP and NIP-Arg
CC contain a potential phosphorylation site comprising G-S-Y; either
CC or both of which S and Y residues have the potential to become
CC phosphorylated. Such phosphorylation may result in the reduction
CC of the inhibitory action of the peptides. A truncated synthetic
CC peptide (AAR56298) with the core sequence of NIP, namely G-S-Y-F-F,
CC was found to inhibit PMN chemotaxis with approx. the same potency
CC as NIP. The invention pertains to peptides which inhibit PMN
CC activation and include within their sequence the sequence AAR56298,
CC or AAR56293, AAR56288; but which contain modifications in those
CC residues S, Y and also G, which form both the target and additional
CC recognitions elements of the phosphorylation site. Such
CC peptide include the second generation sequences
CC in AAR56289-92; AAR56294-97; AAR56299-302; and AAR56304-07. AAR56308 is
CC a peptide with the same proportion of AAs as NIP but in a random
CC sequence. It was produced for use as a control in cellular
CC function studies.

XX SQ Sequence 11 AA;

Query Match 36.9%; Score 31; DB 15; Length 11;

Best Local Similarity 55.6%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 7 RQVEYFGD 15

I: I:|I|

Db 1 REGAYFGD 9

RESULT 10

ABP47080

ID ABP47080 standard; peptide; 13 AA.

XX AC ABP47080;

XX DT 19-AUG-2002 (first entry)

XX DE Human BLYS binding scFv VH CDR3 SEQ ID 3091.

XX KW BLYS; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW common variable immunodeficiency; acquired immunodeficiency syndrome.

XX OS Homo sapiens.

XX PN WO200202641-A1.

XX PD 10-JAN-2002.

XX PF 15-JUN-2001; 2001WO-US19110.

XX PR 16-JUN-2000; 2000US-212210P.

XX PR 17-OCT-2000; 2000US-240816P.

XX PR 21-MAR-2001; 2001US-276248P.

XX PR 16-MAR-2001; 2001US-277379P.

XX PR 25-MAY-2001; 2001US-293499P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.

PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

XX WPI; 2002-114799/15.

XX Antibodies against B lymphocyte stimulating polypeptides, useful for
PT the diagnosis and treatment of cancers and immune disorders -

XX PS Claim 2; Page 3113; 3148pp; English.

XX This invention describes novel antibodies that immunospecifically bind to
CC B lymphocyte stimulator (BLYS) polypeptides. BLYS is a member of the
CC tumour necrosis factor (TNF) super family and induces B cell
CC proliferation and differentiation. The antibodies of the invention have
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
CC antirheumatic and antiAIDS activity and can be used in vaccines to
CC inhibit the expression and activity of BLYS. The antibodies bind to BLYS
CC and so may be used to detect and quantitate the presence of BLYS in
CC biological samples and may be used in this way to diagnose disease
CC associated with aberrant expression of BLYS. They may also be
CC administered to treat diseases associated with aberrant BLYS expression
CC and activity such as cancer, immune, and autoimmune disorders and
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method
CC of the invention.

XX SQ Sequence 13 AA;

Query Match 36.9%; Score 31; DB 23; Length 13;

Best Local Similarity 62.5%; Pred. No. 75;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 RQVEYFGD 14

I: I:|I|

Db 3 RQAYYYTG 10

RESULT 11

AAY86493

ID AAY86493 standard; peptide; 23 AA.

XX AC AAY86493;

XX DT 19-APR-2000 (first entry)

XX DE Human gene 61-encoded protein fragment, SEQ ID NO:408.

XX KW Human; secreted protein; cancer; tumour; developmental abnormality;
KW foetal deficiency; blood disorder; immune system disorder; inflammation;
KW autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KW schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
KW atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KW digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
KW therapy.

XX OS Homo sapiens.

XX PN WO9966041-A1.

XX PD 23-DEC-1999.

XX PF 15-JUN-1999; 99WO-US13418.

XX PR 16-JUN-1998; 98US-0089507.

XX PR 16-JUN-1998; 98US-0089508.

XX PR 16-JUN-1998; 98US-0089509.

XX PR 16-JUN-1998; 98US-0089510.

XX PR 22-JUN-1998; 98US-0090112.

XX PR 22-JUN-1998; 98US-0090113.

XX PA (HUMA-) HUMAN GENOME SCI INC.

PI	Ruben SM, NI J, Rosen CA, Wei Y, Young PE, Florence KA; Soppet DR, Brewer LA, Endress KA, Carter KC, Mucenski M, Lafleur DW, Olsen HS, Shi Y, Moore PA, Komatsoulis G; WPI: 2000-106100/09.	XX	24-OCT-1997; 97US-0063089.	XX	24-OCT-1997; 97US-0063089.
PI		XX	24-OCT-1997; 97US-0063090.	XX	24-OCT-1997; 97US-0063090.
PI		XX	24-OCT-1997; 97US-0063091.	XX	24-OCT-1997; 97US-0063091.
XX		XX	24-OCT-1997; 97US-0063092.	XX	24-OCT-1997; 97US-0063092.
XX		XX	24-OCT-1997; 97US-0063097.	XX	24-OCT-1997; 97US-0063097.
XX		XX	24-OCT-1997; 97US-0063098.	XX	24-OCT-1997; 97US-0063098.
XX		XX	24-OCT-1997; 97US-0063099.	XX	24-OCT-1997; 97US-0063099.
XX		XX	24-OCT-1997; 97US-0063100.	XX	24-OCT-1997; 97US-0063100.
XX		XX	24-OCT-1997; 97US-0063101.	XX	24-OCT-1997; 97US-0063101.
XX		XX	24-OCT-1997; 97US-0063109.	XX	24-OCT-1997; 97US-0063109.
XX		XX	24-OCT-1997; 97US-0063110.	XX	24-OCT-1997; 97US-0063110.
XX		XX	24-OCT-1997; 97US-0063111.	XX	24-OCT-1997; 97US-0063111.
XX		XX	24-OCT-1997; 97US-0063148.	XX	24-OCT-1997; 97US-0063148.
XX		XX	24-OCT-1997; 97US-0063386.	XX	24-OCT-1997; 97US-0063386.
XX		XX	(HUMA-) HUMAN GENOME SCI INC.	XX	(HUMA-) HUMAN GENOME SCI INC.
XX		XX	Brewer LA, Carter KC, Duan DR, Ebner R, Endress KA; Feng P, Florence C, Florence KA, Greene JM, Janat F; Kay H, Lafleur DW, Moore PA, NI J, Olsen HS, Rosen CA; Ruben SM, Shi Y, Soppet DR, Wei Y, Young P; WPI: 1999-303069/25.	XX	Brewer LA, Carter KC, Duan DR, Ebner R, Endress KA; Feng P, Florence C, Florence KA, Greene JM, Janat F; Kay H, Lafleur DW, Moore PA, NI J, Olsen HS, Rosen CA; Ruben SM, Shi Y, Soppet DR, Wei Y, Young P; WPI: 1999-303069/25.
XX		XX	New isolated human genes and the secreted polypeptides they encode	XX	New isolated human genes and the secreted polypeptides they encode
XX		XX	Disclosure: Page 512; 546pp; English.	XX	Disclosure: Page 512; 546pp; English.
XX		XX	The specification describes human secreted proteins. The polynucleotides and their corresponding secreted polypeptides are useful for preventing, treating or ameliorating medical conditions, e.g. by protein or gene therapy. Pathological conditions can also be diagnosed by determining the amount of the polypeptides in a sample or by determining the presence of mutations in the polynucleotides. Specific uses are described for each of the polynucleotides, based on which tissues they are most highly expressed in, and include developing products for the diagnosis or treatment of cancer, tumours, neurodegenerative disorders, developmental abnormalities and fetal deficiencies, blood disorders, leukemias, diseases of the immune system, autoimmune diseases, hepatic and renal disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's and cognitive disorders, schizophrenia, prostate diseases, obesity, disorders involving osteoclasts such as osteoporosis, arthritis or malignancies, diseases of testes, lung or thymus, digestive/endocrine disorders, infections and AIDS. The polypeptides are also useful for identifying their binding partners.	XX	The specification describes human secreted proteins. The polynucleotides and their corresponding secreted polypeptides are useful for preventing, treating or ameliorating medical conditions, e.g. by protein or gene therapy. Pathological conditions can also be diagnosed by determining the amount of the polypeptides in a sample or by determining the presence of mutations in the polynucleotides. Specific uses are described for each of the polynucleotides, based on which tissues they are most highly expressed in, and include developing products for the diagnosis or treatment of cancer, tumours, neurodegenerative disorders, developmental abnormalities and fetal deficiencies, blood disorders, leukemias, diseases of the immune system, autoimmune diseases, hepatic and renal disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's and cognitive disorders, schizophrenia, prostate diseases, obesity, disorders involving osteoclasts such as osteoporosis, arthritis or malignancies, diseases of testes, lung or thymus, digestive/endocrine disorders, infections and AIDS. The polypeptides are also useful for identifying their binding partners.
XX		XX	Query Match 36.9%; Score 31; DB 20; Length 24;	XX	Query Match 36.9%; Score 31; DB 20; Length 24;
XX		XX	Best Local Similarity 60.0%; Pred. No. 1.5e+02;	XX	Best Local Similarity 60.0%; Pred. No. 1.5e+02;
XX		XX	Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	XX	Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX		XX	Sequence 24 AA;	XX	Sequence 24 AA;
XX		XX	2 ERATIRQVEYVFGDF 16	XX	2 ERATIRQVEYVFGDF 16
XX		XX	4 QRAALVLENYKDF 18	XX	4 QRAALVLENYKDF 18
XX		XX	RESULT 12	XX	RESULT 12
XX		XX	AA19735	XX	AA19735
XX		XX	AA19735 standard; Protein: 24 AA.	XX	AA19735 standard; Protein: 24 AA.
XX		XX	AA19735;	XX	AA19735;
XX		XX	14-JUL-1999 (first entry)	XX	14-JUL-1999 (first entry)
XX		XX	DE	XX	DE
XX		XX	SEQ ID NO 453 from W0992243.	XX	SEQ ID NO 453 from W0992243.
XX		XX	Human secreted protein; cancer; tumour; neurodegenerative disorder; developmental abnormality; fetal deficiency; blood disorder; leukemia; immune system disease; autoimmune disease; hepatic disease; renal disease; lymphoma; inflammation; allergy; ischemic shock; Alzheimer's; cognitive disorder; schizophrenia; prostate disease; obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease; lung disease; thymus disease; digestive disorder; endocrine disorder; infection; AIDS.	XX	Human secreted protein; cancer; tumour; neurodegenerative disorder; developmental abnormality; fetal deficiency; blood disorder; leukemia; immune system disease; autoimmune disease; hepatic disease; renal disease; lymphoma; inflammation; allergy; ischemic shock; Alzheimer's; cognitive disorder; schizophrenia; prostate disease; obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease; lung disease; thymus disease; digestive disorder; endocrine disorder; infection; AIDS.
XX		XX	Homo sapiens.	XX	Homo sapiens.
XX		XX	W0992243-Al.	XX	W0992243-Al.
XX		XX	06-MAY-1999.	XX	06-MAY-1999.
XX		XX	23-OCT-1998; 98WO-US22376.	XX	23-OCT-1998; 98WO-US22376.
XX		XX	24-OCT-1997; 97US-0063387.	XX	24-OCT-1997; 97US-0063387.
XX		XX	24-OCT-1997; 97US-0062784.	XX	24-OCT-1997; 97US-0062784.
XX		XX	24-OCT-1997; 97US-0063088.	XX	24-OCT-1997; 97US-0063088.
XX		XX	Neutrophil inhibitor peptide; NIP-Arg; phagocyte inhibitor; degranulation inhibitor; pulmonary inflammation.	XX	Neutrophil inhibitor peptide; NIP-Arg; phagocyte inhibitor; degranulation inhibitor; pulmonary inflammation.
XX		XX	Synthetic.	XX	Synthetic.
XX		XX	Key	XX	Key
XX		XX	Location/Qualifiers	XX	Location/Qualifiers

FT Active-site 3..5
 FT /label- Modified phosphorylation site
 XX WO9414463-A.

XX 07-JUL-1994.

XX 20-DEC-1993; 93WO-US12474.

XX 21-DEC-1992; 92US-0995269.

XX (UABR-) UAB RES FOUND.

XX Cooper JAD;

XX WPI; 1994-234345/28.

XX Novel neutrophil inhibitory peptides - inhibit phagocyte
 PT function, used to reduce pulmonary inflammation

PS Claim 19; Page 15; 97pp; English.

XX Bronchoalveolar lavage was performed on lungs obtd. from heart
 CC transplant donors. At the time of death there was no evidence of
 CC lung infection nor previous lung disease. Inhibitor peptides were
 CC purified from one litre of lavage fluid. Sequence analysis revealed
 CC two peptides - native NIP (AAR56288) and native NIP-Arg (AAR56293). The
 CC sequences exhibited a striking homology to a stretch of AAs in the
 CC sequence of a nucleoprotein prod. by certain influenza A viruses
 CC (AAR56303). Partial homology to sequences within myf-6, ros and neu
 CC oncogene-related proteins was also observed. NIP and NIP-Arg
 CC contain a potential phosphorylation site comprising G-S-Y; either
 CC or both of which S and Y residues have the potential to become
 CC phosphorylated. Such phosphorylation may result in the reduction
 CC of the inhibitory action of the peptides. A truncated synthetic
 CC peptide (AAR56298) with the core sequence of NIP, namely G-S-Y-F-F.
 CC was found to inhibit PMN chemotaxis with approx. the same potency
 CC as NIP. The invention pertains to peptides which inhibit PMN
 CC activation and include within their sequence the sequence AAR56298,
 CC or AAR56293, AAR56288; but which contain modifications in those
 CC residues S, Y and also G, which form both the target and additional
 CC recognition elements of the phosphorylation site. Such
 CC peptides include the second generation sequences
 CC in AAR56289-92; AAR56294-97; AAR56299-302; and AAR56308 is
 CC a peptide with the same proportion of AAs as NIP but in a random
 CC sequence. It was produced for use as a control in cellular
 CC function studies.

XX Sequence 10 AA;

Query Match 35.7%; Score 30; DB 15; Length 10;
 Best Local Similarity 50.0%; Pred. No. 84;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 8 QVEYFEGD 15

Db 1 EASYFEGD 8

RESULT 14
 AAR56307

XX ID AAR56307 standard; Peptide; 11 AA.

XX AC AAR56307;

XX 04-MAR-1995 (first entry)

XX Modified Influenza A nucleoprotein sequence.

DE Neutrophil inhibitor peptide; NIP; phagocyte inhibitor;
 KW degranulation inhibitor; nucleoprotein; Influenza A.

XX Synthetic.

OS

XX Key Location/Qualifiers
 FT Active-site 3..5
 FT /label- modified phosphorylation site
 XX WO9414463-A.

XX 07-JUL-1994.

XX 20-DEC-1993; 93WO-US12474.

XX 21-DEC-1992; 92US-0995269.

XX (UABR-) UAB RES FOUND.

XX Cooper JAD;

XX WPI; 1994-234345/28.

XX Novel neutrophil inhibitory peptides - inhibit phagocyte
 PT function, used to reduce pulmonary inflammation

PS Claim 28; Page 16; 97pp; English.

XX Bronchoalveolar lavage was performed on lungs obtd. from heart
 CC transplant donors. At the time of death there was no evidence of
 CC lung infection nor previous lung disease. Inhibitor peptides were
 CC purified from one litre of lavage fluid. Sequence analysis revealed
 CC two peptides - native NIP (AAR56288) and native NIP-Arg (AAR56293). The
 CC sequences exhibited a striking homology to a stretch of AAs in the
 CC sequence of a nucleoprotein prod. by certain influenza A viruses
 CC (AAR56303). Partial homology to sequences within myf-6, ros and neu
 CC oncogene-related proteins was also observed. NIP and NIP-Arg
 CC contain a potential phosphorylation site comprising G-S-Y; either
 CC or both of which S and Y residues have the potential to become
 CC phosphorylated. Such phosphorylation may result in the reduction
 CC of the inhibitory action of the peptides. A truncated synthetic
 CC peptide (AAR56298) with the core sequence of NIP, namely G-S-Y-F-F.
 CC was found to inhibit PMN chemotaxis with approx. the same potency
 CC as NIP. The invention pertains to peptides which inhibit PMN
 CC activation and include within their sequence the sequence AAR56298,
 CC or AAR56293, AAR56288; but which contain modifications in those
 CC residues S, Y and also G, which form both the target and additional
 CC recognition elements of the phosphorylation site. Such
 CC peptides include the second generation sequences
 CC in AAR56289-92; AAR56294-97; AAR56299-302; and AAR56308 is
 CC a peptide with the same proportion of AAs as NIP but in a random
 CC sequence. It was produced for use as a control in cellular
 CC function studies.

XX Sequence 11 AA;

Query Match 35.7%; Score 30; DB 15; Length 11;
 Best Local Similarity 50.0%; Pred. No. 93;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 8 QVEYFEGD 15

Db 2 EASYFEGD 9

RESULT 15

XX AAE19842

XX ID AAE19842 standard; peptide; 20 AA.

XX AC AAE19842;

XX 18-JUN-2002 (first entry)

XX Human hrpr derived peptide.

XX Human; reaper protein; Rpr; detection; purification; screening;

XX therapy; tumour; cytostatic.

XX OS Homo sapiens.
 XX FN WO200212540-A2.
 XX PD 14-FEB-2002.
 XX PF 08-AUG-2001; 2001WO-US24765.
 XX PR 08-AUG-2000; 2000US-223699P.
 XX PA (UYDU-) UNIV DUKE.
 XX PI Kornbluth SA, Holley C;
 XX DR WPI; 2002-241769/29.
 XX PT New human homologue of Drosophila melanogaster reaper protein (hrpr),
 PT useful for generating antibodies and for screening compounds, which can
 PT inhibit or enhance hrpr activity
 XX
 XX PS Example 1; Page 19; 45pp; English.
 XX

CC The invention relates to human homologue of Drosophila melanogaster
 CC Reaper protein (hrpr) and its corresponding nucleic acid. The hrpr
 CC polypeptides are useful for generating antibodies, which can be used
 CC in detection or purification protocols designed to detect or purify
 CC the polypeptide to which the antibody is directed. These sequences
 CC are also used for screening compounds, which can enhance or inhibit
 CC hrpr and for treating tumours. The hrpr polynucleotides are useful
 CC as a probe or primer. The present sequence is human homologue of
 CC Drosophila melanogaster reaper protein (hrpr) derived peptide.
 XX

SQ Sequence 20 AA;

Query Match 35.7%; Score 30; DB 23; Length 20;
 Best Local Similarity 46.2%; Pred. No. 1.9e+02;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 QERAIQVEYYF 13
 :|: |::| | |
 Db 1 KEKQILROSEVLF 13

Search completed: April 23, 2003, 13:43:02
 Job time : 24.1798 secs

Result No.	Score		Query Match %	Length	DB	ID	Description
	Score	Score					
1	52.5	58.3	390	5	Q8T8V5	Q8T8V5 drosophila	
2	45	50.0	500	16	Q98T70	Q98T70 rhizobium 1	
3	44	48.9	365	10	Q9SV58	Q9SV58 arabidopsis	
4	44	48.9	512	5	Q9VZP8	Q9VZP8 drosophila	
5	43.5	48.3	381	11	Q9CXB9	Q9CXB9 mus musculus	
6	43	47.8	507	3	Q94Q73	Q94Q73 saccharomyc	
7	42	46.7	451	16	Q8YQB2	Q8YQB2 anabaena sp	
8	42	46.7	632	10	O04697	O04697 pisum sativ	
9	42	46.7	642	16	Q9KEU7	Q9KEU7 bacillus ha	
10	42	46.7	935	5	Q9VWZ7	Q9VWZ7 drosophila	
11	41	45.6	308	17	Q8THQ9	Q8THQ9 methanocarc	
12	41	45.6	322	10	Q9C948	Q9C948 arabidopsis	
13	41	45.6	370	16	Q9C4R8	Q9C4R8 caulobacter	
14	41	45.6	415	5	Q8WRU9	Q8WRU9 meloidogyne	
15	41	45.6	633	10	Q9STU9	Q9STU9 arabidopsis	
16	41	45.6	1036	16	O51165	O51165 borrelia bu	


```

DR Pfam: PF00884; Sulfatase: 1.
DR PROSITE: PS00523; SULFATASE_1; 1.
SQ SEQUENCE 512 AA; 58904 MW; AA5B6D8400B6FB3A CRC64;

Query Match 48.3%; Score 44; DB 5; Length 512;
Best Local Similarity 57.1%; Pred. No. 17;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VSKLEASTIRQYY 14
: | | | | | | |
Db 278 ISPLQAQIRQSY 291

RESULT 5
Q9CYB9 PRELIMINARY; PRT; 381 AA.
AC Q9CYB9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Sjogren syndrome antigen B.
GN SSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gotojori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayaishizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK017822; BAB30957.1; -.
DR MGD; MGI:98423; SSB.
DR InterPro: IPR002344; Lupus.La.
DR DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm1.1
DR PRINTS: PRO0302; LOPUSLA.
SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;

Query Match 48.3%; Score 43.5; DB 11; Length 381;
Best Local Similarity 55.6%; Pred. No. 15;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEASTIRQ-EYFQD 17
: | | | | | | |
Db 10 MTALEAKICHQIEYFQD 27

RESULT 6
Q94073 PRELIMINARY; PRT; 507 AA.
AC Q94073;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

```

```

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cystathionine beta-synthase (EC 4.2.1.22).
GN CYS4 OR NHS5 OR YGR155W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X2180-1A;
RA Tezuka H., Mori T., Okumura Y., Kitabatake K., Tsumura Y.;
RT "Cloning of a gene suppression hydropyrogen sulfide production by
RT Saccharomyces cerevisiae and its expression in a brewing yeast.";
RL ASBC Journal 50:130-133(1992).
DR EMBL: D16496; BAA03947.1; -.
DR HSSP; F35520; LJBQ.
DR SGD; S0003387; CYS4.
DR InterPro: IPR001926; B6_enzyme_beta.
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR001216; Cys_synthase.
DR Pfam: PF00571; CBS; 2.
DR Pfam: PF00291; PALP; 1.
DR SMART; SM00116; CBS; 2.
DR TIGRFAMs; TIGR01137; cysta_beta; 1.
DR PROSITE; PS00901; CYS_SYNTHASE; 1.
KW Lyase.
SQ SEQUENCE 507 AA; 56025 MW; 1589165DEB7C8AE4 CRC64;

Query Match 47.8%; Score 43; DB 3; Length 507;
Best Local Similarity 58.8%; Pred. No. 26;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SKLEASTIRQYYFGDA 18
: | | | | | | |
Db 351 SKLEASTTKYADVFGNA 367

RESULT 7
Q8YQB2 PRELIMINARY; PRT; 451 AA.
AC Q8YQB2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE UDP-N-acetylglucosamine pyrophosphorylase.
GN ALR3921.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL: AP003594; BAB75620.1; -.
DR InterPro: IPR001451; Hexapep-transf.
DR InterPro: IPR001825; NTP_transferase.
DR Pfam: PF00132; hexapep; 6.
DR Pfam: PF00483; NTP_transferase; 1.
DR TIGRFAMs; TIGR01173; glmu; 1.
KW Complete proteome.
SQ SEQUENCE 451 AA; 49191 MW; B1DEB3A6F2170FA7 CRC64;

Query Match 46.7%; Score 42; DB 16; Length 451;
Best Local Similarity 53.3%; Pred. No. 36;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```


QY 4 LEASTIRQYFFGDA 18
| : : ||||| ||
Db 187 LEANNAQKEYLTD 201

RESULT 8

Q04697 PRELIMINARY; PRT; 632 AA.
AC 004697;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE DNA-binding protein PD2.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ALASKA; TISSUE=LEAF;
RA Sato N., Kazuno A.A., Ohta N., Ohshima K.;
RT "Identification of a novel family of DNA-binding proteins with two AT-hook motifs from pea."
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: X58740; CAA87292.1; -
KW DNA-binding.
SQ SEQUENCE 632 AA; 69498 MW; 9F744E227CD08717 CRC64;

Query Match 46.7%; Score 42; DB 10; Length 632;
Best Local Similarity 50.0%; Pred. No. 52;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 LEASTIRQYFFGD 17
| : : ||||| ||
Db 161 VENTVSDYITGD 174

RESULT 9

Q9KEU7 PRELIMINARY; PRT; 642 AA.
AC 09KEU7;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ABC transporter (permease).
GN BH0752.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res 28:4317-4331(2000).
DR EMBL: AF001509; BAB0471.1; -
DR InterPro: IPR003838; DUF214.
DR IntraPro: IPR001991; Na/dico_symp.
DR Pfam: PF02687; DUF214; 1.
DR PRINTS: PR00173; EDTNSPORT.
KW Complete proteome.
SQ SEQUENCE 642 AA; 73354 MW; A44515A412FE61E9 CRC64;

Query Match 46.7%; Score 42; DB 16; Length 642;
Best Local Similarity 53.8%; Pred. No. 53;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 EASTIRQYFFGD 17
| : : ||||| ||
Db 472 EKALVLEHYFGD 484

RESULT 10

Q9VWZ7 PRELIMINARY; PRT; 935 AA.
AC 09VWZ7;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE CG6867 protein.
GN CG6867.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananathides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Brokstein P., Brotter P.,
RA Borkova D., Botchan M.R., Bouck J., Brocktein P., Brothier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flesler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwac C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AF003507; AAF48788.1; -
DR FlyBase: FBgn0030887; CG6867.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003112; Olfac_like.
DR Pfam: PF01391; Collagen; 3.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF02191; OLF; 1.
DR ProDom: PD000007; Collagen; 2.
DR SMART: SM00408; IGC2; 2.
DR SMART: SM00284; OLF; 1.

```

KW Collagen; Immunoglobulin domain.
SQ SEQUENCE 935 AA; 102974 MW; 32C26DAC50238B93 CRC64;

Query Match 46.7%; Score 42; DB 5; Length 935;
Best Local Similarity 46.7%; Pred. No. 82;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VSKLEASTIROEYF 15
:|:|:|:|:|:|
Db 827 VAKLDAETLRQYNF 841

RESULT 11
Q8THQ9 PRELIMINARY; PRT; 308 AA.
AC Q8THQ9;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Glycosyltransferase.
GN MA4453.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., DeArelano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Metcalf W.W., Birren B.;
RA "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL; AE01166; AAM07794.1; -.
KW Transferrase; Complete proteome.
SQ SEQUENCE 308 AA; 35336 MW; 01B8F3349B6026E4 CRC64;

Query Match 45.6%; Score 41; DB 17; Length 308;
Best Local Similarity 43.8%; Pred. No. 35;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VSKLEASTIROEYF 16
:|:|:|:|:|:|
Db 105 IGLLDADTVLENYF 120

RESULT 12
Q9C948 PRELIMINARY; PRT; 322 AA.
AC Q9C948;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE Hypothetical 37.2 kDa protein.
GN T7P1.19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;

```

```

RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaya I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana";
RL Nature 408:816-820(2000).
DR EMBL; AC018908; AAG51648.1; -.
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
KW Hypothetical protein.
SQ SEQUENCE 322 AA; 37227 MW; FE951AE12761D1C5 CRC64;

Query Match 45.6%; Score 41; DB 10; Length 322;
Best Local Similarity 64.3%; Pred. No. 37;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VSKLEASTIROEYF 14
:|:|:|:|:|:|
Db 44 VSKLWSSQIRRPY 57

RESULT 13
Q9A4R8 PRELIMINARY; PRT; 370 AA.
ID Q9A4R8;
AC Q9A4R8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Cytochrome c-type biogenesis protein Cych.
GN CC2762.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005942; AAK24726.1; -.
DR TIGR; CC2762; -.
DR InterPro; IPR001440; TPR.
KW Complete proteome.
SQ SEQUENCE 370 AA; 38477 MW; 92B183473D71DE0D CRC64;

Query Match 45.6%; Score 41; DB 16; Length 370;
Best Local Similarity 43.8%; Pred. No. 43;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 3 KLEASTIROEYFGDA 18
:|:|:|:|:|:|

```

Db 220 KVDAADVRRARYLGRA 235

RESULT 14

Q8WRU9 PRELIMINARY; PRT; 415 AA.

AC Q8WRU9; (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Calreticulin.
OS Meloidogyne incognita (southern root-knot nematode).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
ON NCBI_TaxID=6306;
RN [1]
RP SEQUENCE FROM N.A.
RA Jaubert S., Ledger T.N., Plette C., Abad P., Rosso M.N.;
RT "Direct identification of styllet secreted proteins from root-knot
nematodes by a proteomic approach.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF402771; AAL40720.1; -
DR InterPro; IPR001580; Calreticulin.
DR Pfam; PF00262; calreticulin; 1.
DR PRINTS; PR00626; CALRETICULIN.
DR ProDom; PD001866; Calreticulin; 1.
DR PROSITE; PS00803; CALRETICULIN_1; UNKNOWN_1.
DR PROSITE; PS00804; CALRETICULIN_2; UNKNOWN_1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; UNKNOWN_3.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 415 AA; 48596 MW; E23F6B5170EA04DF CRC64;

Query Match 45.6%; Score 41; DB 5; Length 415;
Best Local Similarity 60.0%; Pred. No. 50;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 KLEASTIQEYFVGD 17

Db 114 KLMASTINGEDFHE 128

RESULT 15

Q9STL9 PRELIMINARY; PRT; 633 AA.

AC Q9STL9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 70.8 kDa protein (At3g48390/T29H11.90).
GN T29H11.90.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Choinsne N., Robert C., Brottier P., Wincker P., Cattolico L.,
RA Artiguenave F., Saurin W., Weissenbach J., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,

RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL049659; CAB41159.1; -
DR EMBL; AY074567; AAL67107.1; -
DR InterPro; IPR003891; IF EIF4G_MA3.
DR Pfam; PF02847; MA3; 4.
DR SMART; SM00544; MA3; 4.
KW Hypothetical protein.
SQ SEQUENCE 633 AA; 70751 MW; E7D0FDDBC5E2198 CRC64;

Query Match 45.6%; Score 41; DB 10; Length 633;
Best Local Similarity 53.3%; Pred. No. 80;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 KLEASTIQEYFVGD 17

Db 353 KDAETIIQEYFLSD 367

Search completed: April 23, 2003, 13:32:59
Job time : 29.9101 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:42 ; Search time 4.95506 Seconds
(without alignments)
150.669 Million cell updates/sec

Title: US-09-836-073-18
Perfect score: 90
Sequence: 1 VSKLEASTIRQEEYFGDA 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	79.5	88.3	383	1 LA_AEDAL	Q26457 aedes albop
2	52.5	58.3	390	1 LA_DROME	P40796 drosophila
3	43.5	48.3	404	1 LA_BOVIN	P10881 bos taurus
4	43.5	48.3	408	1 LA_HUMAN	P05455 homo sapien
5	43.5	48.3	415	1 LA_MOUSE	P32067 mus musculu
6	43.5	48.3	415	1 LA_RAT	P38656 rattus norv
7	43.5	48.3	507	1 CBS_YEAST	P32582 saccharomyc
8	40.4	44.4	112	1 YN48_ARCFU	O30321 archaeoglob
9	40.4	44.4	546	1 CK12_YEAST	P32292 saccharomyc
10	39.4	43.3	314	1 PYRB_DEIRA	Q9rvco delnococcus
11	39.4	43.3	431	1 UL61_HCMVA	P16818 human cytom
12	39.4	43.3	533	1 YF59_HUMAN	Q9hcl3 homo sapien
13	39.4	43.3	550	1 SVR_CORGL	P35868 corynebacte
14	39.4	43.3	778	1 S21B_BRARE	Q9w686 brachydanio
15	39.4	43.3	904	1 SVA_SULTO	Q971j4 sulfolobus
16	38.5	42.8	1234	1 CFAH_MOUSE	P06909 mus musculu
17	38.5	42.8	212	1 OMPW_ECOLI	P21364 escherichia
18	38.5	42.8	230	1 VTB8_AGR9	P05357 agrobacteri
19	38.5	42.8	230	1 VTB8_AGR9	P09781 agrobacteri
20	38.5	42.8	540	1 Y968_TREPA	O83934 treponema p
21	38.5	42.8	607	1 HTPG_FUSNN	Q8rgh4 fusobacteri
22	38.5	42.8	771	1 SM3A_HUMAN	Q44563 homo sapien
23	38.5	42.8	772	1 SM3A_CHICK	Q90607 gallus gall
24	38.5	42.8	772	1 SM3A_MOUSE	O08665 mus musculu
25	38.5	42.8	772	1 SM3A_RAT	Q63548 rattus norv
26	37.5	41.7	387	1 SUCC_MYCTU	P71559 mycobacteri
27	37.5	41.7	393	1 SUCC_MYCLE	Q9z5h8 mycobacteri
28	37.5	41.7	428	1 LAA_XENLA	P28048 xenopus lae
29	37.5	41.7	268	1 YIAJ_HAEIN	P44996 haemophilus
30	37.5	41.7	332	1 MYOD_DROME	P22816 drosophila
31	37.5	41.7	339	1 MURG_THEMEA	Q9wy74 thermotoga
32	37.5	41.7	422	1 U183_CAEEL	P34692 caenorhabdi
33	37.5	41.7	427	1 LAB_XENLA	P28049 xenopus lae

34	37	41.1	433	1	SUCB_SCHPO	O94415 schizosacch
35	37	41.1	471	1	YMT1_CAEEL	P98080 caenorhabdi
36	37	41.1	550	1	SVR_MYCSM	Q9x5m0 mycobacteri
37	37	41.1	556	1	SVR_LISIN	Q927c2 listeria in
38	37	41.1	556	1	SVR_LISMO	Q9y493 listeria mo
39	37	41.1	586	1	GGT5_HUMAN	P36269 homo sapien
40	37	41.1	663	1	TRA_BPMU	P07636 bacteriopho
41	37	41.1	775	1	SM3E_HUMAN	O15041 homo sapien
42	37	41.1	775	1	SM3E_MOUSE	P70275 mus musculu
43	37	41.1	875	1	SVV_BORSU	O51680 borrelia bu
44	37	41.1	3137	1	CA36_CHICK	P15989 gallus gall
45	37	41.1	3176	1	CA36_HUMAN	P12111 homo sapien

ALIGNMENTS

RESULT 1
LA_AEDAL STANDARD; PRT; 383 AA.
AC Q26457; 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
OS Aedes albopictus (Forest day mosquito).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Culicoidae; Aedes.
OX NCBI_TaxID=7160;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96135233; PubMed=8551578;
RA Pardigon N., Strauss J.H.;
RT "Mosquito homolog of the La autoantigen binds to Sindbis virus-RNA.";
RL J. Virol. 70:1173-1181(1996).
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTION TERMINATION BY RNA
CC POLYMERASE III. BINDS RNA AND DNA. BINDS TO THE 3' END OF THE
CC MINUS STRAND OF SINDBIS VIRUS RNA. THIS MAY BE SIGNIFICANT FOR
CC SINDBIS VIRUS RNA REPLICATION.
CC -!- SUBCELLULAR LOCATION: NUCLEAR. PRIMARILY NUCLEAR,, BUT SIGNIFICANT
CC AMOUNTS ARE PRESENT IN THE CYTOPLASM.
CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -!- SIMILARITY: TO VERTEBRATE PROTEIN LA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S80954; AAB35931.1; ..
CC InterPro: IPR002344; Lupus.La.
CC Pfam: PF00076; rrm; 1.
CC PRINTS: PR00302; LUPUSLA.
CC SMART: SM00360; RRM; 1.
CC PROSITE: PS0102; RRM; 1.
CC PROSITE: PS00030; RRM_RNP_1; FALSE-NEG.
CC RNA-binding; Nuclear protein; DNA-binding. (RRM).
FT DOMAIN 141 228 RNA-BINDING (RRM).
SQ SEQUENCE 383 AA; 44430 MW; 4E5CC8F21C40F452 CRC64;
Query Match 88.3%; Score 79.5; DB 1; Length 383;
Best Local Similarity 94.7%; Pred. No. 1.8e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 VSKLEASTIRQEEYFGDA 18
|||||
DB 40 VSKLEASTIRQEEYFGDA 58

RESULT 2

ID LA_DROME STANDARD; PRT; 390 AA.

AC P40796; Q24375; Q9VIN2;

AT 01-FEB-1995 (Rel. 31, Created)

CT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE La protein homolog (la ribonucleoprotein) (la autoantigen homolog).

GN LA OR C010922.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_Taxid=7227;

OR [1]

RN "Developmental characterization of a Drosophila RNA-binding protein

RT homologous to the human systemic lupus erythematosus-associated

RT La/SS-B autoantigen.";

RT Mol. Cell. Biol. 14:5123-5129(1994).

RL [2]

RN SEQUENCE FROM N.A.

RP MEDLINE=94309661; PubMed=8035818;

RP Yoo C.J., Wolin S.L.;

RA "La proteins from Drosophila melanogaster and Saccharomyces

RT cerevisiae: a yeast homolog of the La autoantigen is dispensable for

RT growth.";

RT Mol. Cell. Biol. 14:5412-5424(1994).

RL [3]

RN SEQUENCE FROM N.A.

RP STRAIN=Berkeley;

RP MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananthanathan P.G., Scherer S.E., Richards S., Ashburner M., Henderson S.N.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Ball J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,

RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam D.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Lin Y., Lin X.,

RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissensbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

```
CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
CC AND 7-2 RNAS.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
CC C-TERMINAL PART OF THE PROTEIN.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X13698; CAA31986.1; -
CC PIR; S03849; S03849.
CC InterPro; IPR002344; Lupus_La.
CC Pfam; PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS0102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; 1.
CC RNA-binding; Nuclear protein; Phosphorylation.
CC FT DOMAIN 111 187
CC SEQUENCE 404 AA; 46534 MW; 4EE30B5C262AD6A1 CRC64;
CC
CC Query Match 48.3%; Score 43.5; DB 1; Length 404;
CC Best Local Similarity 55.6%; Pred. No. 4.3;
CC Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
CC
QY 1 VSKLEASTIRQ-EYFYGD 17
DB 10 MAALEAKICHQIEYFGD 27
:: ||| | |||||
:: ||| | |||||

RESULT 4
LA_HUMAN STANDARD; PRT; 408 AA.
AC P05455;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lupus La protein (Sjogren syndrome type B antigen (SS-B)) (La
DE ribonucleoprotein) (La autoantigen).
GN SSB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89202037; PubMed=2458131;
RX Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
RT sequences for RNA-binding."
RL Nucleic Acids Res. 17:2233-2244(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89053970; PubMed=3192525;
RX Chambers J.C., Kenan D., Martin B.J., Keene J.D.;
RT "Genomic structure and amino acid sequence domains of the human La
RT autoantigen."
RL J. Biol. Chem. 263:18043-18051(1988).
RN [3]
RP SEQUENCE OF 54-408 FROM N.A.
RX MEDLINE=88199081; PubMed=2452201;
RX Sturgess A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
```

```
RA Coppel R.S.;
RA "Characteristics and epitope mapping of a cloned human autoantigen
RA La."
RL J. Immunol. 140:3212-3218(1988).
RN [4]
RP SEQUENCE OF 54-97 FROM N.A.
RX MEDLINE=85166283; PubMed=3856888;
RA Chambers J.C., Keene J.D.;
RT "Isolation and analysis of cDNA clones expressing human lupus La
RT antigen."
RL Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119(1985).
RN [5]
RP FUNCTION.
RX MEDLINE=89251617; PubMed=2470590;
RX Gottlieb E., Steltz J.A.;
RT "Function of the mammalian La protein: evidence for its action in
RT transcription termination by RNA polymerase III."
RL EMBO J. 8:851-861(1989).
RN [6]
RP PHOSPHORYLATION.
RX MEDLINE=97207017; PubMed=9054510;
RX Pan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Maraie R.J.;
RT "Phosphorylation of the human La antigen on serine 366 can regulate
RT recycling of RNA polymerase III transcription complexes."
RL Cell 88:707-715(1997).
CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
CC AND 7-2 RNAS.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
CC C-TERMINAL PART OF THE PROTEIN.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- DISEASE: SERA FROM PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS
CC OFTEN CONTAIN ANTIBODIES THAT REACT WITH THE NORMAL CELLULAR
CC LA PROTEIN AS IF THESE ANTIGEN WAS FOREIGN.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X13697; CAA31985.1; -
CC EMBL; J04205; AAA51885.1; -
CC PIR; A31888; A31888.
CC PIR; A22956; A22956.
CC PIR; A31273; A31273.
CC PIR; S03848; S03848.
CC PIR; S11013; S11013.
CC Genew; HGNC:11316; SSB.
CC MIM; 109090; -
CC InterPro; IPR002344; Lupus_La.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS0102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; 1.
CC Systemic lupus erythematosus; RNA-binding; Phosphorylation;
KW Nuclear protein.
FT DOMAIN 111 187 RNA-BINDING (RRM).
FT MOD_RES 366 366 PHOSPHORYLATION (BY CK2).
FT SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;
SQ
CC
CC Query Match. 48.3%; Score 43.5; DB 1; Length 408;
CC Best Local Similarity 55.6%; Pred. No. 4.3;
CC Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
```

QY 1 VSKLEASTIRQ-EYFEGD 17
 :: ||| | |||||
 Db 10 MALEAKICHQIEYFEGD 27

RESULT 5

LA_MOUSE STANDARD; PRT; 415 AA.
 AC P32067;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 DE homology.
 GN SSB OR SS-B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93203630; PubMed=8454877;
 RA Topf F., Gordon T., McCluskey J.;
 RT "Characterization of the mouse autoantigen Ia (SS-B). Identification of conserved RNA-binding motifs, a putative ATP binding site and reactivity of recombinant protein with poly(U) and human autoantibodies.";
 RT J. Immunol. 150:3091-3100(1993).
 RL [2]
 RP SEQUENCE OF 1-11 FROM N.A.
 RA Grodz D., Bachmann M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S, AND 7-2 RNAs.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- PTM: PHOSPHORYLATED (BY SIMILARITY).
 CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L00993; AAA39415.1; -;
 CC EMBL; Y07951; CAA69249.1; -;
 CC MGD; MGI:98423; Ssb.
 CC InterPro; IPR002344; Lupus_La.
 CC InterPro; IPR000504; RNA_rec_mot.
 CC Pfam; PF00076; rrm; 1.
 CC PRINTS; PRO0302; LUPUSLA.
 CC SMART; SM00360; RRM; 1.
 CC PROSITE; PS0102; RRM; 1.
 CC PROSITE; PS00030; RRM_RNP_1; 1.
 KW RNA-binding; Nuclear protein; RNA-BINDING (RRM); Phosphorylation.
 FT DOMAIN 111 187
 SQ SEQUENCE 415 AA; 47756 MW; 2D5197692FDC933 CRC64;

Query Match 48.3%; Score 43.5; DB 1; Length 415;
 Best Local Similarity 55.6%; Pred. No. 4.4;
 Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEASTIRQ-EYFEGD 17
 :: ||| | |||||
 Db 10 MTALEAKICHQIEYFEGD 27

RESULT 6

LA_RAT STANDARD; PRT; 415 AA.
 AC P36556;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 DE homology.
 GN SSB OR SS-B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93246255; PubMed=7916708;
 RA Semei I., Troester H., Bartsch H., Schwemmler M., Igloi G.L., Bachmann M.;
 RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La: detection of species-specific variations.";
 RT Gene 126:265-268(1993).
 RL -!- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S, AND 7-2 RNAs.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- PTM: PHOSPHORYLATED (BY SIMILARITY).
 CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X67859; CAA48043.1; -;
 CC PIR; JC1494; JC1494.
 CC InterPro; IPR002344; Lupus_La.
 CC InterPro; IPR000504; RNA_rec_mot.
 CC Pfam; PF00076; rrm; 1.
 CC PRINTS; PRO0302; LUPUSLA.
 CC SMART; SM00360; RRM; 1.
 CC PROSITE; PS0102; RRM; 1.
 CC PROSITE; PS00030; RRM_RNP_1; 1.
 KW RNA-binding; Nuclear protein; RNA-BINDING (RRM); Phosphorylation.
 FT DOMAIN 111 187
 SQ SEQUENCE 415 AA; 47777 MW; 033FD9CC1E475F98 CRC64;

Query Match 48.3%; Score 43.5; DB 1; Length 415;
 Best Local Similarity 55.6%; Pred. No. 4.4;
 Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEASTIRQ-EYFEGD 17
 :: ||| | |||||
 Db 10 MALEAKICHQIEYFEGD 27

RESULT 7

CBS_YEAST STANDARD; PRT; 507 AA.
 AC P32582; Q05177;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cystathionine beta-synthase (EC 4.2.1.22) (Serine sulfhydryase) (Beta-thionase).
 DE (Cys4 OR STR4 OR YGR155W OR G6667).


```
RESULT 9
CK12 YEAST
ID CK12_YEAST STANDARD; PRT; 546 AA.
AC P23292;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 13-JUN-2002 (Rel. 41, Last annotation update)
DE Casein kinase I homolog 2 (EC 2.7.1.-).
GN YCK2 OR CK11 OR YNL154C OR N1755.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RX MEDLINE=92108037; PubMed=1729698;
RA Robinson L.C., Hubbard E.J.A., Graves P.R., de Paoli-Roach A.A.,
RA Roach P.J., Kung C., Haas D.W., Hagedorn C.H., Goebel M.,
RA Culbertson M.R., Carlson M.;
RT "Yeast casein kinase I homologues: an essential gene pair.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:28-32(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92329995; PubMed=1627830;
RA Wang P.-C., Vancura A., Mitcheson T.G.M., Kuret J.;
RA "Two genes in Saccharomyces cerevisiae encode a membrane-bound form
RT of casein kinase-1.";
RL Mol. Biol. Cell 3:275-286(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RX MEDLINE=96287653; PubMed=8686380;
RA Nasr F., Becam A.-M., Herbert C.J.;
RT "The sequence of 36.8 kb from the left arm of chromosome XIV reveals
RT 24 complete open reading frames: 18 correspond to new genes, one of
RT which encodes a protein similar to the human myotonic dystrophy
RT kinase.";
RL Yeast 12:169-175(1996).
CC -1- FUNCTION: CASEIN KINASES ARE OPERATIONALLY DEFINED BY THEIR
CC PREFERENTIAL UTILIZATION OF ACIDIC PROTEINS SUCH AS CASEINS
CC AS SUBSTRATES.
CC -1- SUBCELLULAR LOCATION: Plasma-membrane bound.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CASEIN KINASE I SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M74453; AAA35230.1; -
CC EMBL: X60326; CAA42896.1; -
CC EMBL: X92517; CAA63285.1; -
CC EMBL: Z71430; CAA96041.1; -
CC FIR: A43764; A43764.
CC FIR: S29522; S29522.
CC HSP: P40233; ICSN.
CC SGD: S0005098; YCK2.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR002230; Ser_thr_pkinase.
CC Pfam: PF00669; pkinase; 1.
CC ProDom: PD000001; Euk_pkinase; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Multigene family; Membrane; Prenylation; Lipoprotein.
FT DOMAIN 76 360
PROTEIN KINASE.

FT NP_BIND 82 90 ATP (BY SIMILARITY).
FT BINDING 105 105 ATP (BY SIMILARITY).
FT ACT_SITE 195 195 BY SIMILARITY.
FT LIPID 545 545 GERANYL-GERANYL (POTENTIAL).
FT LIPID 546 546 GERANYL-GERANYL (POTENTIAL).
SQ SEQUENCE 546 AA; 62079 MW; A88D99BC265BF4FA CRC64;

Query Match 44.4%; Score 40; DB 1; Length 546;
Best Local Similarity 62.5%; Pred. No. 25;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 3 KLEAST--IRQEYVFG 16
: | | | | | | | |
DB 124 KILAGTPGIPQEYVFG 139

RESULT 10
PYRB_DEIRA
ID PYRB_DEIRA STANDARD; PRT; 314 AA.
AC Q9VRC0;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aspartate carbamoyltransferase (EC 2.1.3.2) (Aspartate
DE transcarbamylase) (ATCase).
GN PYRB OR DR1109.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Doffat R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-aspartate -> phosphate
CC + N-carbamoyl-L-aspartate.
CC -1- PATHWAY: Pyrimidine biosynthesis; second step.
CC -1- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001961; AAF10682.1; -
CC HSP: P00479; 3CSU.
CC TIGR: DR1109; -
CC InterPro: IPR002029; Asp/Orn_Cotranf.
CC InterPro: IPR002082; Asp_carbmtransf.
CC Pfam: PF00185; OTCace; 1.
CC Pfam: PF02729; OTCace_N; 1.
CC PRINTS: PR00100; AOTCASE.
CC TIGRFAMs: TIGR00670; asp_carb_tr; 1.
CC PROSITE: PS00097; CARBAMOYLTRANSFERASE; 1.
CC Pyrimidine biosynthesis; Transferase; Complete proteome.
KW SEQUENCE 314 AA; 34081 MW; 631875A2AD3EF4B6 CRC64;

Query Match 43.3%; Score 39; DB 1; Length 314;
Best Local Similarity 80.0%; Pred. No. 20;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 4 LEASTIRQY 13
Db 148 LDAYTIRQY 157

RESULT 11
UL61_HCMVA STANDARD; PRT; 431 AA.
AC P16818;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DE 01-FEB-1991 (Rel. 17, Last annotation update)
DE Hypothetical protein UL61.
GN UL61.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchison C.A. III, Kourzides T., Martignetti J.A.,
RA Freddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RT "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169."
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB046779; BABI3385.1; ALT_INIT.
DR HSP; P25490; 1UBD.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; Zf-C2H2; 13.
DR Pfam; PF01352; KRAB; 1.
DR PRINTS; PR00048; ZINC_FINGER.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 13.
DR PROSITE; PS50805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 13.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 13.
KW Hypothetical protein; Transcription regulation; DNA-binding;
KW Zinc-finger; Metal-binding; Nuclear protein; Repeat.
FT DOMAIN 6 78
FT DOMAIN 172 530 ZINC_FINGERS.
FT ZN_FING 172 194 C2H2-TYPE.
FT ZN_FING 200 222 C2H2-TYPE.
FT ZN_FING 228 250 C2H2-TYPE.
FT ZN_FING 256 278 C2H2-TYPE.
FT ZN_FING 284 306 C2H2-TYPE.
FT ZN_FING 312 334 C2H2-TYPE.
FT ZN_FING 340 362 C2H2-TYPE.
FT ZN_FING 368 390 C2H2-TYPE.
FT ZN_FING 396 418 C2H2-TYPE.
FT ZN_FING 424 446 C2H2-TYPE.
FT ZN_FING 452 474 C2H2-TYPE.
FT ZN_FING 480 502 C2H2-TYPE.
FT ZN_FING 508 530 C2H2-TYPE.
SQ SEQUENCE 533 AA; 63463 MW; E40EF5EC22A99F10 CRC64;

Query Match 43.3%; Score 39; DB 1; Length 533;
Best Local Similarity 53.3%; Pred. No. 36;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SKLEASTIRQYFG 16
Db 127 SKIEGEKEQEGYFG 141

RESULT 13
SYR_CORGL STANDARD; PRT; 550 AA.
AC P35868; P41253;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Arginyl-tRNA synthetase (EC 6.1.1.19) (Arginine--tRNA ligase) (ArgRS).
GN ARGS OR CG11179.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
OC Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13059 / AS019;
RX MEDLINE=91186817; PubMed=2082143;
RA Marcel T., Archer J.A.C., Mengin-Lecreulx D., Sinskey A.J.;

```

RT "Nucleotide sequence and organization of the upstream region of the
RT Corynebacterium glutamicum lysa gene.";
RL Mol. Microbiol. 4:1819-1830(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 13869;
RX MEDLINE-94042911; PubMed-8226683;
RA Ogilva J.A., Malumbres M., Eriani G., Pisabarro A., Mateos L.M.,
RA Martin F., Martin J.F.;
RT A gene encoding arginyl-tRNA synthetase is located in the upstream
RT region of the lysa gene in Brevibacterium lactofermentum: regulation
RT of arg-lysA cluster expression by arginine.";
RL J. Bacteriol. 175:7356-7362(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP IDENTIFICATION
RX MEDLINE-93268096; PubMed-8497194;
RA Sharp P.M., Mitchell K.J.;
RT "Corynebacterium glutamicum arginyl-tRNA synthetase.";
RL Mol. Microbiol. 8:200-200(1993).
CC -1- CATALYTIC ACTIVITY: ATP + L-arginine + tRNA(Arg) -> AMP +
CC diphosphate + L-arginyl-tRNA(Arg).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X54740; CAA38537.1; ALT_INIT.
DR EMBL; 221501; CAA79710.1; -.
DR EMBL; AP005277; BAB98572.1; -.
DR PIR; S12227; S12227.
DR PIR; S42850; S42850.
DR InterPro; IPR001278; Arg_tRNA-synt_1c.
DR InterPro; IPR005148; N.
DR InterPro; IPR001412; tRNA-synt_1g; 1.
DR Pfam; PF00750; tRNA-synt_1g; 1.
DR Pfam; PF03485; N-Arg; 1.
DR PRINTS; PR01038; TRNASYNTHARG.
DR TIGRFAMs; TIGR00456; args; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 130 140
FT SITE "HIGH" REGION.
FT SITE 374 378
FT SITE "KMSKS" REGION.
FT BINDING 377 377
FT BINDING ATP (BY SIMILARITY).
FT CONFLICT 355 355
FT CONFLICT G -> D (IN REF. 2).
FT CONFLICT 412 412
FT CONFLICT I -> M (IN REF. 2).
FT CONFLICT 513 513
FT CONFLICT V -> A (IN REF. 2).
FT CONFLICT 540 540
FT CONFLICT H -> R (IN REF. 2).
SQ SEQUENCE 550 AA; 59723 MW; 3AF7248DEE8DC4C1 CRC64;

Query Match 43.3%; Score 39; DB 1; Length 550;
Best Local Similarity 56.2%; Pred. No. 38;
Matches 9; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

QY 4 LEAS--TIQREYFEGD 17
||||| : |||||
DB 154 LEASGAKVTREYFND 169

RESULT 14
SZLB_BRARE

ID SZLB_BRARE STANDARD; PRT; 778 AA.
AC Q9W686;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Semaphorin zlb precursor (Semaphorin 1B) (Sema-zlb).
GN SEMAZ1B OR SEMAZ3AB.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99425174; PubMed-10495275;
RA Roos M., Schachner M., Bernhardt R.R.;
RT "Zebrafish semaphorin zlb inhibits growing motor axons in vivo.";
RL Mech. Dev. 87:103-117(1999).
CC -1- FUNCTION: MIGHT NORMALLY INFLUENCE THE MIDSEGMENTAL PATHWAY CHOICE
CC OF THE VENTRALLY EXTENDING MOTOR AXONS BY CONTRIBUTING TO A
CC REPULSIVE DOMAIN IN THE POSTERIOR SOMITE.
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN RHOMBOMERES THREE AND FIVE, AND
CC IN THE POSTERIOR HALF OF NEWLY FORMED SOMITES WHICH IS AVOIDED BY
CC VENTRALLY EXTENDING MOTOR AXONS.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF083382; AAD28103.1; -.
DR ZFIN; ZDB-GENE-991209-6; sema3ab.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF01437; PSI; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00423; PSI; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 17
FT SIGNAL POTENTIAL.
FT CHAIN 18 778
FT CHAIN SEMAPHORIN 21B.
FT DOMAIN 241 539
FT DOMAIN SEMA.
FT DOMAIN 645 723
FT DOMAIN IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 721 776
FT DOMAIN ARG/LYS-RICH (BASIC).
FT DISULFID 652 716
FT DISULFID BY SIMILARITY.
FT CARBOHYD 54 54
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 127 127
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 593 593
FT CARBOHYD N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 778 AA; 88904 MW; 4D36F4323AE21895 CRC64;

Query Match 43.3%; Score 39; DB 1; Length 778;
Best Local Similarity 38.9%; Pred. No. 55;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 VSKLEASTIROEYFEDA 18
::: ||| : |||
DB 490 ITAMELSTKQQLYLSA 507

RESULT 15
SYA_SULTO

ID SVL_SULTO STANDARD; PRT; 904 AA.
AC Q971J4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine-tRNA ligase) (AlARS).
GN ALAS OR ST1364.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JCM 10545 / 7;
RA MEDLINE=21456156; PubMed=11572479;
RA Kavarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
CC -1- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) -> AMP +
CC diphosphate + L-alanyl-tRNA(Ala).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP000986; BAB66426.1; -
DR InterPro; IPR002318; trna-synt_2c.
DR Pfam; PF01411; trna-synt_2c; 1.
DR TIGRFAMs; TIGR00344; alas; 1.
DR PROSITE; PS50860; AA_TRNA_LIGASE_II_ALA; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 904 AA; 103674 MW; BAABID6F5B08D024 CRC64;

Query Match 43.3%; Score 39; DB 1; Length 904;
Best Local Similarity 47.1%; Fred. No. 66;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 VSKLEASTIRQYFYFGD 17
| | | | | | | | | |
Db 726 VEKIQDGVIRLEYVAGD 742

Search completed: April 23, 2003, 13:28:12
Job time : 5.95506 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:03 ; Search time 10.1124 Seconds
(without alignments)
171.119 Million cell updates/sec

Title: US-09-836-073-18

Perfect score: 90

Sequence: 1 VSKLEASTIRQVEYFGDA 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- PIR_73:*
- 1: pir1.*
 - 2: pir2.*
 - 3: pir3.*
 - 4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	52.5	58.3	390	2 A53773	La/SS-B homolog D-la - fruit fly (Drosophila melanogaster)
2	52.5	58.3	390	2 A53781	ribonucleoprotein
3	44	48.9	365	2 T10213	hypothetical prote
4	43.5	48.3	404	1 S03849	ribonucleoprotein
5	43.5	48.3	408	1 A31888	ribonucleoprotein
6	43.5	48.3	415	1 JCI494	ribonucleoprotein
7	43	47.8	507	2 A48661	cystathionine beta
8	42	46.7	451	2 AB2296	UDP-N-acetylglucos
9	42	46.7	632	2 T06586	DNA-binding protei
10	42	46.7	642	2 H83743	ABC transporter (p
11	41	45.6	322	2 C96636	hypothetical prote
12	41	45.6	370	2 B87591	cytochrome c-type
13	41	45.6	633	2 T06703	hypothetical prote
14	41	45.6	1036	2 D70117	acriflavine resist
15	40	44.4	112	2 D69543	hypothetical prote
16	40	44.4	379	2 A95130	glycogen biosynthe
17	40	44.4	379	2 G88000	required for glyco
18	40	44.4	546	2 S29522	casein kinase I ho
19	40	44.4	835	2 T30030	hypothetical prote
20	40	44.4	876	2 A89944	alanyl-tRNA synthet
21	40	44.4	1201	2 H86434	protein F17F8.21 l
22	39.5	43.9	506	2 H64618	sigma-54 interacti
23	39	43.3	149	2 T08274	hypothetical prote
24	39	43.3	258	2 E71646	hypothetical prote
25	39	43.3	279	2 E37863	hypothetical prote
26	39	43.3	314	2 D75435	aspartate carbamoy
27	39	43.3	426	2 T06086	hypothetical prote
28	39	43.3	431	2 T21594	hypothetical prote
29	39	43.3	431	2 S09824	hypothetical prote

30	39	43.3	517	2 S12227	hypothetical prote
31	39	43.3	550	1 A49936	arginine-tRNA liga
32	39	43.3	801	2 T07617	proteinase TMP - t
33	39	43.3	866	2 E72113	clp proteinase ATP
34	39	43.3	866	2 H86508	clp proteinase ATP
35	39	43.3	872	2 H81556	ATP-dependent clp
36	39	43.3	904	2 AF2304	poly(A) polymerase
37	39	43.3	942	2 S75598	DNA-binding protei
38	38.5	42.8	211	2 G83718	hypothetical prote
39	38.5	42.8	215	2 A59787	probable DNA bindi
40	38.5	42.8	215	2 AG1704	a probable DNA bin
41	38.5	42.8	215	2 AH1333	complement factor
42	38.5	42.8	1234	1 NEMSH	hypothetical prote
43	38	42.2	173	2 T45501	hypothetical prote
44	38	42.2	202	2 T15874	probable outer mem
45	38	42.2	212	2 C85706	

ALIGNMENTS

RESULT 1

A53773

La/SS-B homolog D-la - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000

C:Accession: A53773

R:Bal, C.; Li, Z.; Tolia, P.P.

Mol. Cell. Biol. 14, 5123-5129, 1994

A:Title: Developmental characterization of a Drosophila RNA-binding protein homolog

A:Reference number: A53773; MUID:94309632; PMID:8035794

A:Accession: A53773

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-390 <BAI>

A:Cross-references: GB:U07652; MID:g464019; PIDN:AAA20518.1; PID:g464020

C:Genetics:

A:Gene: FlyBase:La

A:Cross-references: FlyBase:FBgn0011638

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: leucine zipper; RNA binding

Query Match 58.3%; Score 52.5; DB 2; Length 390;

Best Local Similarity 66.7%; Pred. No. 0.17; Mismatches 2; Indels 1; Gaps 1;

Matches 12; Conservative 2;

QY 2 SKLEASTIRQ-EYFQDA 18

DB 48 TKQERAIHQVEYFQDA 65

RESULT 2

A53781

ribonucleoprotein La - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 20-Sep-1999

C:Accession: A53781

R:Yoo, C.J.; Wolin, S.L.

Mol. Cell. Biol. 14, 5412-5424, 1994

A:Title: La proteins from Drosophila melanogaster and Saccharomyces cerevisiae: a y

A:Reference number: A53781; MUID:94309661; PMID:8035818

A:Accession: A53781

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-390 <YOO>

A:Cross-references: GB:I32988; MID:g488469; PID:g488470

C:Genetics:

A:Gene: FlyBase:La

A:Cross-references: FlyBase:FBgn0011638

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: RNA binding

Query Match 58.3%; Score 52.5; DB 2; Length 390;

Best Local Similarity 66.7%; Pred. No. 0.17;
Matches 12; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 2 SKLEASTIRO-EYFQDA 18
: | | : | | | | | | | |
Db 48 TKQERAIROVEYFQDA 65

RESULT 3

Tl0213
hypothetical protein F25G13.200 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999
C:Accession: Tl0213
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke
submitted to the Protein Sequence Database, June 1999
A:Reference number: 216991
A:Accession: Tl0213
A:Molecule type: DNA
A:Residues: 1-365 <BEV>
A:Cross-references: EMBL:AL079349; GSPDB:GN00062; ATSP:F25G13.200
A:Experimental source: cultivar Columbia; BAC clone F25G13
C:Genetics:
A:Gene: ATSP:F25G13.200
A:Map position: 4
A:Introns: 55/2

Query Match 48.9%; Score 44; DB 2; Length 365;

Best Local Similarity 38.9%; Pred. No. 5.6;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VSKLEASTIROEYFQDA 18
: : | | : | | : | |
Db 8 ILRCSSCTMREYFFHDA 25

RESULT 4

S03849
ribonucleoprotein La - bovine
N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C:Species: Bos primigenius taurus (cattle)
C>Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: S03849
R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A:Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences
A:Reference number: S03848; MUID:89202037; PMID:2468131
A:Accession: S03849
A:Molecule type: mRNA
A:Residues: 1-404 <CHA>
A:Cross-references: EMBL:X13698; NID:9755; PIDN:CAA31986.1; PID:g756
A:Note: part of this sequence was confirmed by protein sequencing
C:Comment: This protein associates with a variety of small RNA molecules, most of which
ay act as a transcription termination factor.
C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
C:Keywords: blocked amino end; phosphoprotein; RNA binding
F:112-178/Domain: ribonucleoprotein repeat homology <RNA>
F:113-118/Region: RNA-binding RNP2 motif
F:151-158/Region: RNA-binding RNP1 motif
F:228-404/Domain: phosphorylated #status predicted <PHY>

Query Match 48.3%; Score 43.5; DB 1; Length 404;

Best Local Similarity 55.6%; Pred. No. 7.8;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

Qy 1 VSKLEASTIRO-EYFQD 17
: : | | : | | : | |
Db 10 MAALAKTCHQIEYFQD 27

RESULT 5

A31888
ribonucleoprotein La - human

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome ant
C:Species: Homo sapiens (man)
C>Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: A31888; S03848; A22956; A61051; S11013; I55553; I70205; I70206; A31273
R:Chambers, J.C.; Kenan, D.; Martin, B.J.; Keene, J.D.
J. Biol. Chem. 263, 18043-18051, 1988
A:Title: Genomic structure and amino acid sequence domains of the human La autoantig
A:Reference number: A31888; MUID:89053970; PMID:3192525
A:Accession: A31888
A:Molecule type: mRNA
A:Residues: 1-408 <CHA>
A:Cross-references: GB:J04205; NID:g178686; PIDN:AAA51885.1; PID:g178687
R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.
Nucleic Acids Res. 17, 2233-2244, 1989
A:Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequen
A:Reference number: S03848; MUID:89202037; PMID:2468131
A:Accession: S03848
A:Molecule type: mRNA
A:Residues: 1-408 <CH2>
A:Cross-references: EMBL:X13697; NID:g36414; PIDN:CAA31985.1; PID:g36415
R:Chambers, J.C.; Keene, J.D.
Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985
A:Title: Isolation and analysis of cDNA clones expressing human lupus La antigen.
A:Reference number: A22956; MUID:85166283; PMID:3856886
A:Accession: A22956
A:Molecule type: mRNA
A:Residues: 45-97, 'LK' <CH3>
A:Cross-references: GB:J04205
A:Note: This sequence has been revised in reference A31888
R:Nyman, U.; Ringertz, N.R.; Pettersson, I.
Immunol. Lett. 22, 65-72, 1989
A:Title: Demonstration of an amino terminal La epitope recognized by human anti-La s
A:Reference number: A61051; MUID:89379261; PMID:2476379
A:Accession: A61051
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-19, 'E', '21-47 <NYM>
R:Sturgess, A.D.; Peterson, M.G.; McNeillage, L.J.; Whittingham, S.; Coppel, R.L.
J. Immunol. 140, 3212-3218, 1988
A:Title: Characteristics and epitope mapping of a cloned human autoantigen La.
A:Reference number: S11013; MUID:88199081; PMID:2452201
A:Accession: S11013
A:Molecule type: mRNA
A:Residues: 'E', 55-287, 'V', 289-408 <STU>
A:Cross-references: EMBL:M20328; NID:g337456; PIDN:AAA36577.1; PID:g337457
R:Kohsaka, H.; Yamamoto, K.; Fujii, H.; Miura, H.; Miyasaka, N.; Nishioaka, K.; Miyar
J. Clin. Invest. 85, 1566-1574, 1990
A:Title: Fine epitope mapping the human SS-B/La protein: Identification of a distinc
A:Reference number: I55553; MUID:90237237; PMID:1692037
A:Accession: I55553
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 81-107 <RES>
A:Cross-references: GB:M35261; NID:g338491; PIDN:AAA36652.1; PID:g338495
A:Accession: I70205
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 174-224 <RE2>
A:Cross-references: GB:M35263; NID:g338492; PIDN:AAA36653.1; PID:g338496
A:Accession: I70206
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 279-342 <RE3>
A:Cross-references: GB:M35262; NID:g338493; PIDN:AAA36654.1; PID:g338497
C:Comment: This protein associates with a variety of small RNA molecules, most of w
ay act as a transcription termination factor.
C:Genetics:
A:Gene: GDB:SSB
A:Cross-references: GDB:125359; OMIM:109090
A:Map position: 2
A:Introns: 22/3; 57/2; 115/3; 185/2; 209/2; 223/3; 264/3; 380/2
C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
C:Keywords: phosphoprotein; RNA binding

F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:1228-408/Domain: phosphorylated #status experimental <PHY>

Query Match 48.3%; Score 43.5; DB 1; Length 408;
 Best Local Similarity 55.6%; Pred. No. 7, 9;
 Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEASTIRQ-EYFVGD 17
 :: ||| | |||||
 Db 10 MAALEAKICHQIEYFVGD 27

RESULT 6

JCI1494

ribonucleoprotein La - rat

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B

C:Species: Rattus norvegicus (Norway rat)

C:Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: JCI1494; S25145

R:Sensei, I.; Troester, H.; Bartsch, H.; Schwemmler, M.; Igloi, G.L.; Bachmann, M.

Gene 126, 265-268, 1993

A:Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection of s

A:Reference number: JCI1494; MUID:93246255; PMID:7916708

A:Accession: JCI1494

A:Molecule type: mRNA

A:Residues: 1-415 <SEM>

A:Cross-references: GB:X67859; NID:g55778; PIDN:CAAM48043.1; PID:g55779

A:Experimental source: liver

C:Comment: This protein associates with a variety of small RNA molecules, most of which

ay act as a transcription termination factor.

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: phosphoprotein; RNA binding

F:112-178/Domain: ribonucleoprotein repeat homology <RRM>

F:113-118/Region: RNA-binding RNP2 motif

F:151-158/Region: RNA-binding RNP1 motif

F:1227-415/Domain: phosphorylated #status predicted <PHY>

Query Match 48.3%; Score 43.5; DB 1; Length 415;
 Best Local Similarity 55.6%; Pred. No. 8;
 Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEASTIRQ-EYFVGD 17
 :: ||| | |||||
 Db 10 MAALEAKICHQIEYFVGD 27

RESULT 7

A48661

cystathionine beta-synthase (EC 4.2.1.22) - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein G667; protein YGR155W

C:Species: Saccharomyces cerevisiae

C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 21-Jul-2000

C:Accession: A48661; S48505; S42686; S60445; B55760; S64464; S33202

R:Cherest, H.; Thomas, D.; Surdian-Kerjan, Y.

J. Bacteriol. 175, 5366-5374, 1993

A:Title: Cysteine biosynthesis in Saccharomyces cerevisiae occurs through the transsulfu

A:Reference number: A48661; MUID:93374830; PMID:8366024

A:Accession: A48661

A:Molecule type: DNA

A:Residues: 1-507 <CHE>

A:Cross-references: GB:X72922; NID:g296134; PIDN:CAA51426.1; PID:g296135

R:Ono, B.I.; Inoue, T.; Kijima, K.; Matsuda, A.; Negishi, K.; Shinoda, S.

submitted to the EMBL data Library, June 1993

A:Description: Identification of the structural gene of cystathionine beta-synthase in s

A:Reference number: S48505

A:Accession: S48505

A:Molecule type: DNA

A:Residues: 1-292, 'T', 294-507 <ONO>

A:Cross-references: EMBL:D16502; NID:g391939; PIDN:BAAO3952.1; PID:g416161

R:Ono, B.I.; Kijima, K.; Inoue, T.; Miyoshi, S.I.; Matsuda, A.; Shinoda, S.

Yeast 10, 333-339, 1994

A:Title: Purification and properties of Saccharomyces cerevisiae cystathionine beta

A:Reference number: S42686; MUID:94287710; PMID:8017103

A:Accession: S42686

A:Molecule type: protein

A:Residues: 2-11 <ON2>

R:Skala, J.; Nawrocki, A.; Goffeau, A.

Yeast 11, 1421-1427, 1995

A:Title: The sequence of a 27 kb segment on the right arm of chromosome VII from Sa

A:Reference number: S60435; MUID:96158062; PMID:8585325

A:Accession: S60445

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-507 <SKA>

A:Cross-references: EMBL:X85807; NID:g1045249; PIDN:CAAS9812.1; PID:g1045260

A:Note: the nucleotide sequence was submitted to the EMBL data Library, March 1995

R:Kruger, W.D.; Cox, D.R.

Proc. Natl. Acad. Sci. U.S.A. 91, 6614-6618, 1994

A:Title: A yeast system for expression of human cystathionine beta-synthase: struct

A:Reference number: A55760; MUID:94294429; PMID:8022826

A:Accession: B55760

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptua

A:Molecule type: DNA

A:Residues: 1-61, 63-128, 'Y', 130-406, 'Y', 408-435, 'VE', 438-440, 'Y', 442-480, 'E', 482-50

A:Cross-references: GB:L14578

R:Van Dyck, L.; Skala, J.; de Wergifosse, P.; Purnelle, B.; Talla, E.; Nawrocki, A.

submitted to the Protein Sequence Database, May 1996

A:Reference number: S64428

A:Accession: S64464

A:Molecule type: DNA

A:Residues: 1-507 <VAN>

A:Cross-references: EMBL:Z72940; NID:g1323262; PIDN:CAA97169.1; PID:g1323263; MIPS:

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:CYS4; STR4

A:Cross-references: SGD:S0003387; MIPS:YGR155W

A:Map position: 7R

C:Superfamily: cystathionine beta-synthase; CBS homology

C:Keywords: carbon-oxygen lyase; cysteine biosynthesis; homotetramer; hydro-lyase;

F:2-507/Product: cystathionine beta-synthase #status experimental <MAT>

F:376-424/Domain: CBS homology <CBS>

F:53/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 47.8%; Score 43; DB 2; Length 507;
 Best Local Similarity 58.8%; Pred. No. 12;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SKLEASTIRQYFVGD 18
 ||||| : ||:
 Db 351 SKLEASTTRYADVFGNA 367

RESULT 8

AB2296

UDP-N-acetylglucosamine pyrophosphorylase [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Title: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AB2296

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iri

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabo

DNA Res. 8, 205-213, 2001

A:Title: Complete genomic sequence of the filamentous nitrogen-fixing Cyanobacteriu

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AB2296

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-451 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAW5620.1; PID:g17133055; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr3921

C:Superfamily: N-acetylglucosamine-1-phosphate uridylyltransferase

Query Match 46.7%; Score 42; DB 2; Length 451;
Best Local Similarity 53.3%; Pred. No. 17;
Matches 8; Conservative 3; Mismatches 4; Indels 0;

QY 4 LEASTIRQYFQDA 18
||||: ||||| ||
DB 187 LEANNAQREYLTDA 201

RESULT 9

T06586
DNA-binding protein PD2 - garden pea
C:Species: Pisum sativum (garden pea)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999
C:Accession: T06586
R:Sato, N.; Kazuno, A.A.; Ohta, N.; Ohshima, K.
submitted to the EMBL Data Library, June 1996
A:Description: Identification of a novel family of DNA-binding proteins with two AT-hook
A:Reference number: 215774
A:Accession: T06586
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-632 <SAT>
A:Cross-references: EMBL:X98740; NID:e995229; PIDN:CAA67292.1; PID:e275185
A:Experimental source: cv. Alaska

Query Match 46.7%; Score 42; DB 2; Length 632;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 LEASTIRQYFQCD 17
||||: ||||| ||
DB 161 VEATVWSDEYFGD 174

RESULT 10

H83743
ABC transporter (permease) BH0752 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: H83743
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: H83743
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-642 <STO>
A:Cross-references: GB:AP001509; GB:BA000004; NID:gl0173176; PIDN:BA04471.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH0752

Query Match 46.7%; Score 42; DB 2; Length 642;
Best Local Similarity 53.8%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 EASTIRQYFQGD 17
|:|:| |||||
DB 472 EKALVHEFYFGD 484

RESULT 11

C96636
hypothetical protein T7Pl.19 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96636
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kin
C.A.; Li, J.H.; Liu, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzi
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tal
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: C96636
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-322 <STO>
A:Cross-references: GB:AE005173; NID:96751694; PIDN:AAF27677.1; GSPDB:GN00141
C:Genetics:
A:Gene: T7Pl.19
A:Map position: 1

Query Match 45.6%; Score 41; DB 2; Length 322;
Best Local Similarity 64.3%; Pred. No. 17;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VSKLEASTIRQYFY 14
||||: ||||| ||
DB 44 VSKLWSSQIRRPYY 57

RESULT 12

H87591
cytochrome c-type biogenesis protein Cych [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: H87591
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelber
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; K
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser,
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete genome sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87591
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-370 <STO>
A:Cross-references: GB:AE005673; NID:gl3424358; PIDN:AAK24726.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2762

Query Match 45.6%; Score 41; DB 2; Length 370;
Best Local Similarity 43.8%; Pred. No. 20;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 KLEASTIRQYFQDA 18
|:|:| |||||
DB 220 KVDAADVRARYVLGRA 235

RESULT 13

T06703
hypothetical protein T29H11.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 22-Oct-1999
C:Accession: T06703
R:Quetier, F.; Choisme, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; A
submitted to the Protein Sequence Database, April 1999
A:Reference number: 215793
A:Accession: T06703
A:Molecule type: DNA
A:Residues: 1-633 <QUE>
A:Cross-references: EMBL:AL049659; GSPDB:GN000061; ATSP:T29H11.90
A:Experimental source: cultivar Columbia; BAC clone T29H11
C:Genetics:
A:Gene: ATSP:T29H11.90
A:Map position: 3
A:Introns: 29/1

Query Match 45.68; Score 41; DB 2; Length 633;
Best Local Similarity 53.3%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 KLEASTIROEYFGD 17
I : I I I I I I I
DB 353 KKDAETIIQYFLSD 367

RESULT 14
D70117
acriflavine resistance protein (acrB) homolog - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 17-Mar-2000
C:Accession: D70117
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
i Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: D70117
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1036 <KLE>
A:Cross-references: GB:AE000783; TIGR:BB0140
A:Experimental source: strain B31
C:Superfamily: cation efflux system membrane protein czca

Query Match 45.68; Score 41; DB 2; Length 1036;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 VSKLEASTIROEYFG 16
I I I I I I I I I
DB 82 VSSKESSTVSLFHYG 97

RESULT 15
D69543
hypothetical protein AF2348 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: D69543
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: D69543
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-112 <KLE>
A:Cross-references: GB:AE001114; GB:AE000782; NID:g2689437; PIDN:AAB91315.1; PID:g265074
C:Superfamily: Archaeoglobus fulgidus hypothetical protein AF2348

Query Match 44.48; Score 40; DB 2; Length 112;
Best Local Similarity 47.1%; Pred. No. 7.7;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 SKLEASTIROEYFGDA 18
I : I I I I I I I
DB 3 SRSRKSIVLQDIYRGDS 19

Search completed: April 23, 2003, 13:34:42
Job time : 11.1124 secs

CC autoantibody and may therefore be used as vaccines.

XX Sequence 21 AA;

Query Match 71.6%; Score 73; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 ICHQIEYFGDF 18

Db 1 ICHQIEYFGDF 12

RESULT 8

ABB65316
ID ABB65316 standard; Protein; 390 AA.

XX AC ABB65316;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 22740.

XX KW Drosophila: developmental biology; cell signalling; insecticide;
pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI: 2001-656860/75.

XX DR N-PSDB; ABL09419.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions

XX PS Disclosure; SEQ ID NO 22740; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
sequences (ABL01840-ABL16175) and the encoded proteins
(ABB5737-ABB7202).

XX CC The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 390 AA;

Query Match 55.9%; Score 57; DB 22; Length 390;
Best Local Similarity 62.5%; Pred. No. 0.46;

Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QQAKICHOIEYFGD 17

Db 49 KQRAIRQVEYFGD 64

RESULT 9

AAG47714

XX ID AAG47714 standard; Protein; 913 AA.

XX AC AAG47714;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 60168.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0123788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 28-APR-1999; 99US-0130891.

XX PR 30-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 04-MAY-1999; 99US-0132407.

XX PR 05-MAY-1999; 99US-0132484.

XX PR 06-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 07-MAY-1999; 99US-0132487.

XX PR 11-MAY-1999; 99US-0132863.

XX PR 14-MAY-1999; 99US-0134256.

XX PR 14-MAY-1999; 99US-0134218.

XX PR 14-MAY-1999; 99US-0134219.

XX PR 14-MAY-1999; 99US-0134221.

XX PR 14-MAY-1999; 99US-0134370.

XX PR 18-MAY-1999; 99US-0134768.

XX PR 19-MAY-1999; 99US-0134941.

XX PR 20-MAY-1999; 99US-0135124.

XX PR 21-MAY-1999; 99US-0135353.

XX PR 24-MAY-1999; 99US-0135629.

XX PR 25-MAY-1999; 99US-0136021.

XX PR 27-MAY-1999; 99US-0136392.

XX PR 28-MAY-1999; 99US-0136782.

XX PR 01-JUN-1999; 99US-0137222.

XX PR 03-JUN-1999; 99US-0137528.

XX PR 04-JUN-1999; 99US-0137502.

XX PR 07-JUN-1999; 99US-0137724.

XX PR 08-JUN-1999; 99US-0138094.

XX PR 10-JUN-1999; 99US-0138540.

XX PR 10-JUN-1999; 99US-0138847.

XX PR 14-JUN-1999; 99US-0139119.

XX PR 16-JUN-1999; 99US-0139452.

XX PR 16-JUN-1999; 99US-0139453.

XX PR 17-JUN-1999; 99US-0139492.

XX PR 18-JUN-1999; 99US-0139454.

XX PR 18-JUN-1999; 99US-0139455.

XX PR 18-JUN-1999; 99US-0139456.

XX PR 18-JUN-1999; 99US-0139457.

XX PR 18-JUN-1999; 99US-0139458.

XX	Arabidopsis thaliana protein fragment SEQ ID NO: 60167.	
DE		99US-0139817.
XX		99US-0139899.
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX		99US-0140353.
XX		99US-0140354.
OS	Arabidopsis thaliana.	
XX		99US-0140695.
XX		99US-0140823.
XX		99US-0140991.
XX		99US-0141287.
XX		99US-0141842.
XX		99US-0142154.
XX		99US-0142055.
XX		99US-0142390.
XX		99US-0142803.
XX		99US-0142920.
XX		99US-0142977.
XX		99US-0143542.
XX		99US-0143624.
XX		99US-0144005.
XX		99US-0144085.
XX		99US-0144331.
XX		99US-0144332.
XX		99US-0144333.
XX		99US-0144334.
XX		99US-0144335.
XX		99US-0144352.
XX		99US-0144632.
XX		99US-0144884.
XX		99US-0144884.
XX		99US-0145086.
XX		99US-0145088.
XX		99US-0145085.
XX		99US-0145087.
XX		99US-0145089.
XX		99US-0145192.
XX		99US-0145145.
XX		99US-0145218.
XX		99US-0145224.
XX		99US-0145276.
XX		99US-0145913.
XX		99US-0145918.
XX		99US-0145919.
XX		99US-0145951.
XX		99US-0146386.
XX		99US-0146388.
XX		99US-0146389.
XX		99US-0147038.
XX		99US-0147204.
XX		99US-0147302.
XX		99US-0147192.
XX		99US-0147260.
XX		99US-0147303.
XX		99US-0147416.
XX		99US-0147493.
XX		99US-0147935.
XX		99US-0148171.
XX		99US-0148319.
XX		99US-0148341.
XX		99US-0148565.
XX		99US-0148684.
XX		99US-0149368.
XX		99US-0149175.
XX		99US-0149426.
XX		99US-0149722.
XX		99US-0149723.
XX		99US-0149929.
XX		99US-0149902.
XX		99US-0149930.
XX		99US-0150566.
XX		99US-0150884.
XX		99US-0151063.
XX		99US-0151066.

PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0158596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 47.1%; Score 48; DB 21; Length 993;
Best Local Similarity 66.7%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIEYFEGD 17
||:|||||
Db 196 KIVNQVEYFSD 207

RESULT 12

ABB59107
ID ABB59107 standard; Protein; 1900 AA.

XX AC ABB59107;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 4113.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI: 2001-656860/75.
 DR N-PSDB; ABL03210.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 4113; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1900 AA;
 Query Match 44.18; Score 45; DB 22; Length 1900;
 Best Local Similarity 50.08; Pred. No. 2.3e+02;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 OY 1 QQQEAKICHQIEY 14
 I I I : I I : I I
 DB 1275 QQRELDCHREEFY 1288
 RESULT 13
 ABG27058
 ID ABG27058 standard; Protein; 135 AA.
 XX AC ABG27058;
 AC ABG27058;
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #27049.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 FA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR N-PSDB; AAS91245.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess

PT biodiversity -
 XX
 PS Claim 20; SEQ ID No 57417; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 135 AA;
 Query Match 43.18; Score 44; DB 22; Length 135;
 Best Local Similarity 70.08; Pred. No. 19;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 4 EAKICHQIEY 13
 I : I I I I I I I
 DB 9 EKRICHQIEF 18
 RESULT 14
 ABB9645
 ID ABB9645 standard; Protein; 224 AA.
 XX AC ABB9645;
 AC ABB9645;
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 2021.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-US16450.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 FA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI: 2002-122018/16.
 DR N-PSDB; ABL90054.
 XX
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative

PT disorders -

XX Claim 11; SEQ ID NO 2021; 2081pp + Sequence Listing; English.

XX The invention relates to novel genes (ABL89449-ABL90853) and proteins (ABB89040-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 224 AA;

Query Match 43.1%; Score 44; DB 23; Length 224;

Best Local Similarity 70.0%; Pred. No. 33;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHQIEY 13

DB 112 EXRICHQIEF 121

RESULT 15

ABB59213

ID ABB59213 standard; Protein; 277 AA.

XX

AC ABB59213;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 4431.

XX

KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

KX

XX Drosophila melanogaster.

OS

XX WO200171042-A2.

PN

XX 27-SEP-2001.

PD

XX

PF 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-191637P.

PR

XX 11-JUL-2000; 2000US-0614150.

XX

PA (PEKE) PE CORP NY.

XX

PI Venter JC, Adams M, Li PWD, Myers EW;

XX

DR WPI; 2001-656860/75.

DR

XX N-PSDB; ABL03316.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

PT

XX Disclosure; SEQ ID NO 4431; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 277 AA;

Query Match 43.1%; Score 44; DB 22; Length 277;

Best Local Similarity 50.0%; Pred. No. 42;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 QOQEKIKCHQIEY 14

DB 187 QMRQEKLCQPEY 200

Search completed: April 23, 2003, 13:27:11

Job time : 29.5169 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:34:48 ; Search time 20.0225 Seconds

(without alignments)
185.234 Million cell updates/sec

Title: US-09-836-073-4

Perfect score: 102

Sequence: 1 QQEAKICHOIEYFQDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 9297

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	32.4	24	5 Q9BM09	Q9bm09 spongilla 1
2	31	30.4	24	3 Q07140	Q07140 saccharomyc
3	29	28.4	16	2 Q9R963	Q9r963 helicobacte
4	28	27.5	11	4 Q9UC46	Q9uc46 homo sapien
5	28	27.5	24	2 Q05616	Q05616 staphylococ
6	27	26.5	20	10 Q9FUY3	Q9fuy3 zea mays (m
7	27	26.5	23	4 Q9UC18	Q9uc18 homo sapien
8	26	25.5	8	4 Q9BY5	Q9by5 homo sapien
9	26	25.5	8	6 Q9BFA0	Q9bfa0 macaca mula
10	26	25.5	8	6 Q9BF99	Q9bf99 hylobates c
11	26	25.5	24	7 Q9TNS7	Q9tns7 homo sapien
12	26	25.5	24	7 Q9TNS6	Q9tns6 homo sapien
13	25	24.5	25	8 Q9TGB8	Q9tgb8 alnus crisp
14	25	24.5	25	8 Q9TGB7	Q9tgb7 alnus gluti
15	25	24.5	25	8 Q9TGB6	Q9tgb6 alnus marit
16	25	24.5	25	8 Q9TGB5	Q9tgb5 betula alle

17	25	24.5	25	8 Q9TGB4	Q9tgb4 betula glan
18	25	24.5	25	8 Q9TGB3	Q9tgb3 betula papy
19	25	24.5	25	8 Q9TGB2	Q9tgb2 betula verr
20	25	24.5	25	8 Q9TGB1	Q9tgb1 betula pube
21	25	24.5	25	8 Q9TGB0	Q9tgb0 corylus ave
22	25	24.5	25	8 Q9TGA9	Q9tga9 corylus col
23	25	24.5	25	8 Q9TGA8	Q9tga8 corylus cor
24	25	24.5	25	8 Q9TGA7	Q9tga7 ostrya virg
25	25	24.5	25	8 Q9TGA6	Q9tga6 quercus rub
26	25	24.5	25	8 Q9TGA3	Q9tga3 carpinus ca
27	24	23.5	12	10 Q9M433	Q9m433 lotus japon
28	24	23.5	13	8 Q9THR8	Q9thr8 bryopsis sp
29	24	23.5	13	12 Q9E1V4	Q9e1v4 hepatitis b
30	24	23.5	13	12 Q9E1V3	Q9e1v3 hepatitis b
31	24	23.5	13	12 Q9E1V2	Q9e1v2 hepatitis b
32	24	23.5	13	12 Q9E1V1	Q9e1v1 hepatitis b
33	24	23.5	13	12 Q9E1V0	Q9e1v0 hepatitis b
34	24	23.5	13	12 Q9E1U9	Q9e1u9 hepatitis b
35	24	23.5	13	12 Q9E1U8	Q9e1u8 hepatitis b
36	24	23.5	13	12 Q9E1U6	Q9e1u6 hepatitis b
37	24	23.5	13	12 Q9E1U5	Q9e1u5 hepatitis b
38	24	23.5	13	12 Q9E1U4	Q9e1u4 hepatitis b
39	24	23.5	13	12 Q9E1U3	Q9e1u3 hepatitis b
40	24	23.5	13	12 Q9E1U2	Q9e1u2 hepatitis b
41	24	23.5	13	12 Q9E1U1	Q9e1u1 hepatitis b
42	24	23.5	13	12 Q9E1U0	Q9e1u0 hepatitis b
43	24	23.5	13	12 Q9E1T9	Q9e1t9 hepatitis b
44	24	23.5	13	12 Q9E1T8	Q9e1t8 hepatitis b
45	24	23.5	21	10 Q41496	Q41496 solanum tub

ALIGNMENTS

RESULT 1

Q9BM09 PRELIMINARY; PRT: 24 AA.

AC Q9BM09; TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE Gypsy-like reverse transcriptase (fragment).

OS Spongilla lacustris (Freshwater sponge).

OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;

OC Haplosclerida; Spongillidae; Spongillia.

OX NCBI_TaxID=6055;

RN [1]

RP SEQUENCE FROM N.A.

RC TRANSPOSON-GRT-G7 RETROTRANSPOSON;

RX MEDLINE=20570504; PubMed=11121049;

RA Arkhipova I., Meselson M.;

RT "Transposable elements in sexual and ancient asexual taxa.";

RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).

DR EMBL; AY013997; AAG59969.1; -

KW RNA-directed DNA polymerase.

FT NON_TER 1

FT NON_TER 24

SQ SEQUENCE 24 AA; 2754 MW; B588512EDE7A251A CRC64;

Query Match 32.4%; Score 33; DB 5; Length 24;
Best Local Similarity 44.4%; Pred. No. 1.5e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIEYF 15

Db 15 VCHGLEFVF 23

RESULT 2

Q07140 PRELIMINARY; PRT: 24 AA.

ID Q07140

AC Q07140;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Class II transposable element TY1-17 5' end (Fragment).
 GN YCLO19W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ND 40-4C;
 RX MEDLINE=85242090; PubMed=2989787;
 RA Fulton A.M., Mellor J., Dobson M.J., Chester J., Warrington J.R.,
 RA Kingsman K.J., Oliver S.G., de la Paz P., Wilson W., Kingsman A.J.,
 RA Kingsman S.M.;
 RT "Variants within the yeast Ty sequence family encode a class of
 RT structurally conserved proteins.";
 RL Nucleic Acids Res 13:4097-4111(1985).
 DR EMBL; X02546; CAA26399.1; -;
 DR SGD; S0000524; YCLO19W.
 FT NON_TER 24 24
 SQ SEQUENCE 24 AA; 2642 MW; A6B0BDF527A58B3 CRC64;
 Query Match 30.4%; Score 31; DB 3; Length 24;
 Best Local Similarity 50.0%; Pred. No. 3.1e+02;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 QOQEKICHOIE 12
 ID Q9R963 PRELIMINARY; PRT; 16 AA.
 AC Q9R963;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE PepC (Fragment).
 GN PEPC.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=F31;
 RX MEDLINE=98453456; PubMed=9780260;
 RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,
 RA Kuriyama M.;
 RT "Full-length sequence analysis of the vacA gene from cytotoxic and
 RT noncytotoxic Helicobacter pylori.";
 RL J. Infect. Dis. 178:1391-1398(1998).
 DR EMBL; AF049623; AAD04263.1; -;
 FT NON_TER 1 1
 FT NON_TER 16 16
 SQ SEQUENCE 16 AA; 1938 MW; 3C33D403AEE61428 CRC64;
 Query Match 28.4%; Score 29; DB 2; Length 16;
 Best Local Similarity 55.6%; Pred. No. 4.4e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 10 QIEYFGDF 18
 ID Q9UC46 PRELIMINARY; PRT; 11 AA.
 AC Q9UC46;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE Neutrophil inhibitor peptide, NIP-POLYMORPHONUCLEAR neutrophil
 DE inhibitor peptide.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96326114; PubMed=8703476;
 RA Cooper J.A.Jr., Culbreth R.R.;
 RT "Characterization of a neutrophil inhibitor peptide harvested from
 RT human bronchial lavage: homology to influenza A nucleoprotein.";
 RL Am. J. Respir. Cell Mol. Biol. 15:207-215(1996).
 SQ SEQUENCE 11 AA; 1262 MW; 951A1C3279C9DB45 CRC64;
 Query Match 27.5%; Score 28; DB 4; Length 11;
 Best Local Similarity 80.0%; Pred. No. 4.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 13 YVFGD 17
 ID Q05616 PRELIMINARY; PRT; 24 AA.
 AC Q05616;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE 3-dehydroquinase synthase (EC 4.6.1.3) (Fragment).
 GN AROB.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8325-4;
 RX MEDLINE=93381456; PubMed=83711108;
 RA O'Connell C.M., Pattee P., Foster T.J.;
 RT "Sequence and mapping of the araC gene of Staphylococcus aureus 8325-4.";
 RL J. Gen. Microbiol. 139:1449-1460(1993).
 CC -|- CATALYTIC ACTIVITY: 7-PHOSPHO-3-DEOXY-ARABINO-HEPTULOSONATE - 3-
 CC DEHYDROQUINATE + ORTHOPHOSPHATE.
 CC -|- PATHWAY: SECOND STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE
 CC AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
 CC -|- SUBUNIT: MONOMER (BY SIMILARITY).
 DR EMBL; L05004; AAA71896.1; -;
 KW Aromatic amino acid biosynthesis; Lyase.
 FT NON_TER 1 1
 FT NON_TER 24 24
 SQ SEQUENCE 24 AA; 2857 MW; 1C4C8D20F7A4856E CRC64;
 Query Match 27.5%; Score 28; DB 2; Length 24;
 Best Local Similarity 50.0%; Pred. No. 9.7e+02;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 8 CHOIEYVF 15
 ID Q9FUY3 PRELIMINARY; PRT; 20 AA.
 AC Q9FUY3;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```

DE MADS-box protein (Fragment).
GN TKZ-1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, INBRED A188; TISSUE=LEAF;
RA Emishev V.Y., Zbrodina M.V., Karyagina A.S., Naroditsky B.S.,
RA Khavkin E.E.;
RT "An unusual k-box sequence of a maize MADS-box gene.";
RL Russ. J. Plant Physiol. 47:555-557(2000).
DR EMBL; AF276682; AAG10438.1; -.
FT NON_TER 1
FT NON_TER 20
FT NON_TER 20
SQ SEQUENCE 20 AA; 2477 MW; 0DFD32A46903C4D0 CRC64;

Query Match 26.5%; Score 27; DB 10; Length 20;
Best Local Similarity 41.7%; Pred. No. 1.2e+03;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QQEAKICHQIE 12
Db 2 QQESPKLRFNQI 13

RESULT 7
Q9UCL8 PRELIMINARY; PRT; 23 AA.
AC Q9UCL8;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 14, Last annotation update)
DE GLYCOSAMINOGLYCAN-BEARING polypeptide (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=93098810; PubMed=1463459;
RA Bonnet F., Perin J.P., Maillet P., Jolles P., Alliel P.M.;
RT "Characterization of a human seminal plasma glycosaminoglycan-bearing
RT polypeptide.";
RL Biochem. J. 288:565-569(1992).
SQ SEQUENCE 23 AA; 2334 MW; E469397E6B2EA794 CRC64;

Query Match 26.5%; Score 27; DB 4; Length 23;
Best Local Similarity 35.3%; Pred. No. 1.4e+03;
Matches 6; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 2 QQAKICHQIEYFGDF 18
Db 1 KQAVSCSEEQETXGDF 17

RESULT 8
Q9BYYS PRELIMINARY; PRT; 8 AA.
AC Q9BYYS;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CAMP responsive element moderator (Fragment).
GN CREM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011664; AAG47575.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 8 AA; 1006 MW; DF02C331EEAB572A CRC64;

Query Match 25.5%; Score 26; DB 4; Length 8;
Best Local Similarity 60.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 8 CHQIE 12
Db 4 CHKVE 8

RESULT 9
Q9BFAO PRELIMINARY; PRT; 8 AA.
AC Q9BFAO;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CAMP responsive element moderator (fragment).
GN CREM.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011661; AAG47572.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 8 AA; 1006 MW; DF02C331EEAB572A CRC64;

Query Match 25.5%; Score 26; DB 6; Length 8;
Best Local Similarity 60.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 8 CHQIE 12
Db 4 CHKVE 8

RESULT 10
Q9BF99 PRELIMINARY; PRT; 8 AA.
AC Q9BF99;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CAMP responsive element moderator (Fragment).
GN CREM.
OS Hylobates concolor (crested gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=29089;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011662; AAG47573.1; -.

```

```
FT NON_TER 1 1
SQ SEQUENCE 8 AA; 1006 MW; DF02C331EAB572A CRC64;

Query Match 25.5%; Score 26; DB 6; Length 8;
Best Local Similarity 60.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHOIE 12
   ||::|
Db 4 CHRVE 8

RESULT 11
Q9TNS7 PRELIMINARY; PRT; 24 AA.
AC Q9TNS7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE Human leucocyte antigen beta chain DR molecule HLA-DRB1 (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93107319; PubMed=1469092;
RA Weyand C.M., Hicok K.C., Hunder G.G., Goronzy J.J.;
RT "The HLA-DRB1 locus as a genetic component in giant cell arteritis.
RT Mapping of a disease-linked sequence motif to the antigen binding site
RT of the HLA-DR molecule."
RL J. Clin. Invest. 90:2355-2361(1992).
KW MHC.
SQ SEQUENCE 24 AA; 3042 MW; 46FF753670C7A760 CRC64;

Query Match 25.5%; Score 26; DB 7; Length 24;
Best Local Similarity 20.0%; Pred. No. 2.1e+03;
Matches 2; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIEYF 15
   ::|::|::|
Db 3 QVRKHEFDYF 12

RESULT 12
Q9TNS6 PRELIMINARY; PRT; 24 AA.
AC Q9TNS6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE Human leucocyte antigen beta chain DR molecule HLA-DRB1 (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93107319; PubMed=1469092;
RA Weyand C.M., Hicok K.C., Hunder G.G., Goronzy J.J.;
RT "The HLA-DRB1 locus as a genetic component in giant cell arteritis.
RT Mapping of a disease-linked sequence motif to the antigen binding site
RT of the HLA-DR molecule."
RL J. Clin. Invest. 90:2355-2361(1992).
KW MHC.
SQ SEQUENCE 24 AA; 3112 MW; 5C4F753667F7A760 CRC64;

Query Match 25.5%; Score 26; DB 7; Length 24;
Best Local Similarity 20.0%; Pred. No. 2.1e+03;
Matches 2; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIEYF 15
   ::|::|::|
Db 3 QVRKHEFDYF 12

RESULT 13
Q9TGB8 PRELIMINARY; PRT; 25 AA.
AC Q9TGB8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Alnus crispa.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=3518;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RT angiosperms and homology to nad5 intron 1."
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL: AF080075; AAD50062.1;
DR InterPro: IPR001351; Ribosomal_S3.
DR Pfam: PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 24.5%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFYG 16
   ::|::|::|
Db 4 DYYG 8

RESULT 14
Q9TGB7 PRELIMINARY; PRT; 25 AA.
AC Q9TGB7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Alnus glutinosa (Alder).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=3517;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RT angiosperms and homology to nad5 intron 1."
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL: AF080076; AAD50063.1;
DR InterPro: IPR001351; Ribosomal_S3.
DR Pfam: PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 24.5%; Score 25; DB 8; Length 25;
```

Best Local Similarity 60.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
:|:|
Db 4 DYYG 8

RESULT 15

Q9TGB6 PRELIMINARY; PRT; 25 AA.
AC Q9TGB6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Alnus maritima.
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=21015;
RN [1]
RP SEQUENCE FROM N.A. PubMed-10331271;
RX MEDLINE=99261656; PubMed-10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RT angiosperms and homology to nad5 intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080077; AAD50064.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 24.5%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
:|:|
Db 4 DYYG 8

Search completed: April 23, 2003, 13:47:15
Job time : 21.0225 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:33:13 ; Search time 4.65169 Seconds
(without alignments)
160.495 Million cell updates/sec

Title: US-09-836-073-4

Perfect score: 102

Sequence: 1 QQQEAKICHQIEYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1520

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	26.5	19	1 HBB2_UROHA	P18992 uromastix h
2	24	23.5	11	1 CXLI_CONMR	P58807 conus marino
3	24	23.5	13	1 CXLI_CONMR	P58810 conus marino
4	24	23.5	16	1 MLE_SQUAC	P01207 squallus aca
5	24	23.5	18	1 MLE_SCYCA	P01206 scyllorhinu
6	24	23.5	25	1 CXOB_CONMA	P05485 conus magus
7	22	21.6	21	1 BRX_ATRBI	P80163 atractaspis
8	22	21.6	21	1 SRPD_ATREN	P13211 atractaspis
9	22	21.6	25	1 ANDT_ANDAU	P56684 androctonus
10	21	20.6	10	1 GON2_CHEPR	P80678 chelyosoma
11	21	20.6	12	1 TIN2_HOPTI	P82652 hoplobatrach
12	21	20.6	17	1 TPIS_PIPNS	P81666 pinus pinas
13	21	20.6	21	1 REV_HV2D2	P15830 human immun
14	21	20.6	22	1 DP20_CAEBR	P51558 caenorhabdi
15	21	20.6	22	1 FUC1_RAT	P80347 rattus norv
16	21	20.6	22	1 FUC1_RAT	P80347 rattus norv
17	21	20.6	23	1 NUO5_SOLTU	P80262 solanum tub
18	21	20.6	23	1 PRO3_DACGL	P18690 dactylis gl
19	21	20.6	25	1 SNBP_RAT	P80968 rattus norv
20	20	19.6	10	1 TRNB_RANRI	P29135 rana ridibu
21	20	19.6	14	1 CXAL_CONCN	P56973 conus conso
22	20	19.6	15	1 PC20_BRANA	P81096 brassica na
23	20	19.6	16	1 CXA2_CONMA	P56636 conus magus
24	20	19.6	18	1 OBP_LYMDI	P34173 lymantria d
25	20	19.6	18	1 SFAN_HELAN	P81098 helianthus
26	20	19.6	20	1 BIP_PHAVU	P80089 phaseolus v
27	20	19.6	20	1 COG4_CHIOP	P34156 chionocete
28	20	19.6	20	1 FIBB_FELCA	P14469 felis silve
29	20	19.6	20	1 TL22_SPIOL	P82796 spinacia ol
30	20	19.6	20	1 YQAH_KLEAE	P56506 klebsiella
31	20	19.6	21	1 NDK_CANAL	Q9ur66 candida alb
32	20	19.6	22	1 LP1_TRIWA	P58930 trimeresuru
33	20	19.6	22	1 LP2_TRIWA	P58930 trimeresuru

34 20 19.6 23 1 XYC1_ACIGB P46365 acinetobact
35 19.5 19.1 20 1 COXN_THUOB P80980 thunnus obe
36 19.5 19.1 23 1 UDP_LACCA P19662 lactobacill
37 19 18.6 8 1 AL17_CARMA P81820 carcinus ma
38 19 18.6 9 1 AL11_CARMA P81814 carcinus ma
39 19 18.6 10 1 TRNB_ONCMY P28500 oncorhynchu
40 19 18.6 11 1 TIN4_HOPTI P82654 hoplobatrach
41 19 18.6 12 1 NO40_SESRO O24369 sesbania ro
42 19 18.6 12 1 TIN3_HOPTI P82653 hoplobatrach
43 19 18.6 14 1 ADF_TENMO P82965 tenebrio mo
44 19 18.6 19 1 FIBB_VULVU P14482 vulpes vulp
45 19 18.6 20 1 CS21_STRTR P81621 streptococc

ALIGNMENTS

RESULT 1
HBB2_UROHA
ID HBB2_UROHA STANDARD: PRT; 19 AA.
AC P18992;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hemoglobin beta-2 chain (Fragment).
OS Uromastix hardwickii (Indian spiny-tailed lizard).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Uromastycinae;
OC Uromastix.
OX NCBI_TaxID=40250;
RN [1]
RP SEQUENCE.
RX MEDLINE=84029159; PubMed=6628672;
RA Nagvi S., Zaidi Z.H., von Bahr-Lindstroem H., Carlquist M.,
RA Joernvall H.;
RT "Characterization of hemoglobin from the lizard Uromastix
hardwickii.";
RL FEBS Lett. 162:290-295(1983).
CC - FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
VARIOUS PERIPHERAL TISSUES.
CC - SUBUNIT: HETEROETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC - TISSUE SPECIFICITY: RED BLOOD CELLS.
CC - SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR; A05305; A05305.
DR InterPro; IPR000971; Globin.
DR PROSITE; PS01033; GLOBIN; PARTIAL.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT NON_TER 1
FT NON_TER 19
SQ SEQUENCE 19 AA; 1914 MW; C40AD8EA30019057 CRC64;
Query Match 26.5%; Score 27; DB 1; Length 19;
Best Local Similarity 80.0%; Pred No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 14 YFGDF 18
Db 1 FFGDF 5
RESULT 2
CXLI_CONMR STANDARD: PRT; 11 AA.
ID CXLI_CONMR
AC P58807;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda-conotoxin CmrVIA.
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]

```

RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE-Venom;
RX MEDLINE-20564325; PubMed-10988292;
RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Sew K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus.";
RL J. Biol. Chem. 275:39516-39522(2000).
CC -1- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- MASS SPECTROMETRY: MW-1237.93; MWERR-0.21; METHOD-Electrospray.
CC -1- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KW Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 2 11
FT DISULFID 3 8
FT MOD_RES 10 10 HYDROXYLATION.
SQ SEQUENCE 11 AA; 1226 MW; 277AAC60B7232B58 CRC64;

Query Match 23.5%; Score 24; DB 1; Length 11;
Best Local Similarity 75.0%; Pred. No. 4.1e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 KICH 9
Db 6 KLCH 9

RESULT 3
CXLA_CONMR STANDARD; PRT; 13 AA.
ID CXLA_CONMR
AC P36810;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Lambda/conotoxin MrIB (Chi-MrIB).
OS Conus marmoreus (Marble Cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]
RP SEQUENCE, SYNTHESIS, AND STRUCTURE BY NMR.
RC TISSUE-Venom;
RX MEDLINE-21419681; PubMed-11528421;
RA Sharpe I.A., Gehrmann J., Loughnan M.L., Thomas L., Adams D.A.,
RA Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.,
RA Lewis R.J.;
RT "Two new classes of conopeptides inhibit the alpha1-adrenoceptor and
RT noradrenaline transporter.";
RL Nat. Neurosci. 4:902-907(2001).
CC -1- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- PTM: Exists in two forms, due to cis-trans isomerization at His-
CC 11-Hyp-12.
CC -1- MASS SPECTROMETRY: MW-1393.52; METHOD-Electrospray.
CC -1- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
DR PDB; IIEQ; 03-APR-02.
KW Neurotoxin; Toxin; Hydroxylation; 3D-structure.
FT DISULFID 4 13
FT DISULFID 5 10
FT MOD_RES 12 12 HYDROXYLATION.
SQ SEQUENCE 13 AA; 1382 MW; 277AAC376EAD2B58 CRC64;

Query Match 23.5%; Score 24; DB 1; Length 13;
Best Local Similarity 75.0%; Pred. No. 4.9e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 KICH 9
Db 8 KLCH 11

```

```

RESULT 4
MLB_SQUAC STANDARD; PRT; 16 AA.
ID MLB_SQUAC
AC P01207;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanotropin beta.
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE.
RX MEDLINE-75127390; PubMed-4375978;
RA Bennett H.P.J., Lowry P.J., McMartin C., Scott A.P.;
RT "Structural studies of alpha-melanocyte-stimulating hormone and a
RT novel beta-melanocyte-stimulating hormone from the neurointermediate
RT lobe of the pituitary of the dogfish Squalus acanthias.";
RL Biochem. J. 141:439-444(1974).
CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
DR PIR; A01471; MTDPBS.
KW Hormone.
SQ SEQUENCE 16 AA; 1939 MW; 993AF43C8A640A0E CRC64;

Query Match 23.5%; Score 24; DB 1; Length 16;
Best Local Similarity 57.1%; Pred. No. 6e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 12 EYFQDF 18
Db 4 DYKFGHF 10

RESULT 5
MLB_SCYCA STANDARD; PRT; 18 AA.
ID MLB_SCYCA
AC P01206;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanotropin beta (Beta-MSH).
OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyliorhinidae; Scyliorhinus.
OX NCBI_TaxID=7830;
RN [1]
RP SEQUENCE.
RX MEDLINE-75113445; PubMed-4452470;
RA Love R.M., Pickering B.T.;
RT "A beta-MSH in the pituitary gland of the spotted dogfish
RT (Scyliorhinus canicula): isolation and structure.";
RL Gen. Comp. Endocrinol. 24:398-404(1974).
CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
DR PIR; A01470; MTDPBS.
KW Hormone.
SQ SEQUENCE 18 AA; 2195 MW; 353C93A30C314DF4 CRC64;

Query Match 23.5%; Score 24; DB 1; Length 18;
Best Local Similarity 44.4%; Pred. No. 6.7e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 QIEYFQDF 18
Db 2 ZIBYKMGHF 10

RESULT 6
CXOB_CONMA STANDARD; PRT; 25 AA.
ID CXOB_CONMA
AC P05405;

```


15-JUN-2002 (Rel. 41, Last annotation update)
 DE Androctonin.
 OS Androctonus australis hector (Sahara scorpion).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Buthidae; Buthidae; Androctonus.
 OX NCBI_TaxID=70175;
 RN [1]
 RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
 RC TISSUE-Hemolymph; PubMed=8939880;
 RX MEDLINE=97094646; PubMed=8939880;
 RA Enret-Sabatier L., Loew D., Goyffon M., Fehlbaum P., Hoffmann J.A.,
 RA van Dorsselaer A., Bulet P.;
 RT "Characterization of novel cysteine-rich antimicrobial peptides from
 RT scorpion blood";
 RL J. Biol. Chem. 271:29537-29544(1996).
 RN [2]
 RP SYNTHESIS OF D-AMINO ACID ENANTIOMER, AND CHARACTERIZATION.
 RX MEDLINE=20115101; PubMed=10642525;
 RA Hetru C., Letellier L., Oren Z., Hoffmann J.A., Shai Y.;
 RT "Androctonin, a hydrophilic disulphide-bridged non-haemolytic
 RT anti-microbial peptide: a plausible mode of action";
 RL Biochem. J. 345:653-664(2000).
 RN [3]
 RP STRUCTURE BY NMR.
 RX MEDLINE=20025109; PubMed=10563585;
 RA Mandard N., Sy D., Maufrais C., Bonmatin J.M., Bulet P., Hetru C.,
 RA Vovelle F.;
 RT "Androctonin, a novel antimicrobial peptide from scorpion Androctonus
 RT australis: solution structure and molecular dynamics simulations in
 RT the presence of a lipid monolayer";
 RL J. Biomol. Struct. Dyn. 17:367-380(1999).
 CC -1- FUNCTION: ACTIVE AGAINST BOTH BACTERIA (GRAM-POSITIVE AND GRAM-
 CC NEGATIVE) AND FILAMENTOUS FUNGI. ACTS ON THE MEMBRANE OF THE
 CC BACTERIAL CELLS. IT DESTABILIZE A MEMBRANE BY MODIFYING ITS
 CC PROPERTIES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MASS SPECTROMETRY: MW=3076.7; METHOD=Electrospray.
 DR PDB; 1C26; 12-JAN-00.
 KW Antibiotic; Fungicide; 3D-structure.
 FT DISULFID 4 20
 FT DISULFID 10 16
 SQ SEQUENCE 25 AA; 3081 MW; D4183D6EA5AA50AA CRC64;
 Query Match 21.6%; Score 22; DB 1; Length 25;
 Best Local Similarity 50.0%; Pred. No. 2e+03;
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 7 ICHQIE 12
 : : :
 DB 3 VCRQIK 8
 RESULT 10
 GON2_CHEPR STANDARD; PRT; 10 AA.
 ID GON2_CHEPR STANDARD; PRT; 10 AA.
 AC P80678;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GNRH-II)
 DE (Luliberin II).
 OS Chelyosoma productum.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 OC Phlebobranchia; Corellidae; Chelyosoma.
 OX NCBI_TaxID=71177;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96413669; PubMed=8816823;
 RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
 RA Powell M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
 RT "Two new forms of gonadotropin-releasing hormone in a protochordate
 RT and the evolutionary implications";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).

-1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
 CC FOLLICLE-STIMULATING HORMONES.
 CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
 CC THE GONODUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
 CC THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
 CC -1- MASS SPECTROMETRY: MW=1117.52; METHOD=MALDI.
 CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
 CC InterPro: IPR002012; GNRH.
 DR PROSITE; PS00473; GNRH; 1.
 KW Hormone; Amidation.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT DISULFID 6 6 INTERCHAIN.
 FT DISULFID 10 10 AMIDATION (BY SIMILARITY).
 SQ SEQUENCE 10 AA; 1135 MW; 284B38D1EEB735A3 CRC64;
 Query Match 20.6%; Score 21; DB 1; Length 10;
 Best Local Similarity 42.9%; Pred. No. 1.1e+03;
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 3 QEAKICH 9
 : : :
 DB 1 QHWSLCH 7
 RESULT 11
 TIN2_HOPII STANDARD; PRT; 12 AA.
 ID TIN2_HOPII STANDARD; PRT; 12 AA.
 AC P82652;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tigerin-2.
 OS Hoplobatrachus tigerinus (Indian bull frog) (Rana tigerina).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranidae;
 OC Hoplobatrachus.
 OX NCBI_TaxID=103373;
 RN [1]
 RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND DISULFIDE BONDS.
 RC TISSUE=Skin;
 RX PubMed=11031261;
 RA Purna Sai K., Jagannadham M.V., Vairamani M., Raju N.P.,
 RA Devi A.S., Nagaraj R., Sitaram N.;
 RT "Tigerinins: novel antimicrobial peptides from the Indian frog Rana
 RT tigerina";
 RL J. Biol. Chem. 276:2701-2707(2001).
 CC -1- FUNCTION: ANTIBACTERIAL ACTIVITY AGAINST B.SUBTILIS, E.COLI,
 CC S.AUREUS, M.LUTEUS, P.PUTIDA AND S.CEREVISIAE.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: SKIN.
 CC -1- MASS SPECTROMETRY: MW=1368; METHOD=MALDI.
 KW Amphibian skin; Antibiotic; Amidation.
 FT MOD_RES 12 11
 FT DISULFID 3 12
 SQ SEQUENCE 12 AA; 1369 MW; C2627D8EFD37605D CRC64;
 Query Match 20.6%; Score 21; DB 1; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 ICH 9
 : : :
 DB 10 ICH 12
 RESULT 12
 TPIS_PINPS STANDARD; PRT; 17 AA.
 ID TPIS_PINPS STANDARD; PRT; 17 AA.
 AC P81666;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update)
 Triosephosphate isomerase, cytosolic (EC 5.3.1.1) (ITM) (Fragments).
 Pinus pinaster (Maritime pine).
 Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Coniferales; Pinaceae; Pinus.
 NCBI_TaxID=71647;
 [1]
 SEQUENCE.
 RQ TISSUE=Needle;
 RX MEDLINE=9274088; PubMed=10344291;
 RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
 RA Elgerio J.-M., Plomion C.;
 RT "Separation and characterization of needle and xylem maritime pine
 proteins";
 RL Electrophoresis 20:1098-1108(1999).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate -> glycerone
 phosphate.
 CC -1- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- INDUCTION: BY WATER STRESS.
 CC -1- MISCELLANEOUS: IN PLANTS, THERE ARE TWO TYPES OF TPIS, CYTOSOLIC
 AND PLASTID.
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN
 (SPOT N139) IS: 5.9, ITS MW IS: 24 kDa.
 CC -1- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
 DR InterPro: IPR000652; Triophos.ismrse.
 DR PROSITE: PS00171; ITM; PARTIAL.
 KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
 KW Pentose shunt.
 FT NON_TER 1 1
 FT NON_CONS 9 10
 FT NON_TER 17 17
 SQ SEQUENCE 17 AA; 2053 MW; 183DB41757AF13CB CRC64;
 Query Match 20.6%; Score 21; DB 1; Length 17;
 Best Local Similarity 20.0%; Pred. No. 1.9e+03;
 Matches 2; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 QY 7 ICHQIEYVFG 16
 :|: :|
 DB 4 VCYEQLFFVG 13
 RESULT 13
 REV_HV2D2
 ID REV_HV2D2 STANDARD; PRT; 21 AA.
 AC P15830;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE REV-Protein (Anti-repression transactivator protein) (ART/TRS)
 DE (Fragment).
 GS Human immunodeficiency virus type 2 (isolate D205.7) (HIV-2).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11716;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90081881; PubMed=2594088;
 RA Dietrich U., Adamski M., Kreutz R., Seipp A., Kuehnelt H.,
 RA Ruebsamen-Waigmann H.;
 RL "A highly divergent HIV-2-related isolate";
 RL Nature 342:948-950(1989).
 CC -1- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE
 NUCLEATE REPRESSION OF GAG AND ENV PRODUCTION.
 CC -1- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleoli.
 CC -1- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED
 BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X61240; -, NOT_ANNOTATED_CDS.
 DR S08441; S08441.
 DR HIV; X16109; REV\$2D205.
 DR InterPro: IPR000625; REV_protein.
 DR Pfam: PF00424; REV; 1.
 KW Transcription regulation; AIDS; Phosphorylation; Nuclear protein.
 FT NON_TER 21 21
 SQ SEQUENCE 21 AA; 2503 MW; E620E225CC5BFF24 CRC64;
 Query Match 20.6%; Score 21; DB 1; Length 21;
 Best Local Similarity 37.5%; Pred. No. 2.4e+03;
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 2 QOEAKICH 9
 :|: :|
 DB 9 QKELRLH 16
 RESULT 14
 DP20_CAEBR
 ID DP20_CAEBR STANDARD; PRT; 22 AA.
 AC P51558;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein dpy-20 (Fragment).
 GN DPY-20.
 OS Caenorhabditis briggsae.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6238;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95287857; PubMed=7770042;
 RA Clark D.V., Suleman D.S., Beckenbach K.A., Gilchrist E.J.,
 RA Baillie D.L.;
 RT "Molecular cloning and characterization of the dpy-20 gene of
 Caenorhabditis elegans";
 RL Mol. Gen. Genet. 247:367-378(1995).
 CC -1- FUNCTION: INVOLVED IN CUTICLE FUNCTION AND IS ESSENTIAL FOR
 CC NORMAL MORPHOLOGICAL DEVELOPMENT (BY SIMILARITY).
 DR TRANSFAC; T04319; -.
 KW Cuticle.
 FT NON_TER 22 22
 SQ SEQUENCE 22 AA; 2529 MW; A11FB717AD367F69 CRC64;
 Query Match 20.6%; Score 21; DB 1; Length 22;
 Best Local Similarity 42.9%; Pred. No. 2.5e+03;
 Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QOQEAKI 7
 :|: :|
 DB 15 QQQQSOL 21
 RESULT 15
 FUC1_RAT
 ID FUC1_RAT STANDARD; PRT; 22 AA.
 AC P80347;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Fucinin 1 (Fucosyltransferase inhibitor 1) (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]

RP SEQUENCE.
 RC STRAIN-Sprague-Dawley; TISSUE=Small intestine mucosa;
 RX MEDLINE=94357231; PubMed=8076650;
 RA Ruggiero-Lopez D., Manioc C., Geourjon C., Louisot P., Martin A.;
 RT "Purification and partial amino acid sequence of fucinin, an
 endogenous inhibitor of fucosyltransferase activities.";
 RL Eur. J. Biochem. 224:47-55(1994).
 CC -1- FUNCTION: HAS A ROLE IN THE PHYSIOLOGICAL REGULATION OF
 FUCOSYLATION PROCESSES.
 CC -1- SUBUNIT: OLIGOMER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: SOME, WITH HUMAN SET/PHAPII PROTEIN.
 FT NON_TER 22 22
 SQ SEQUENCE 22 AA; 2393 MW; 0A12574A68A8E8A9 CRC64;
 Query Match 20.6%; Score 21; DB 1; Length 22;
 Best Local Similarity 80.0%; Pred. No. 2.5e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QQQEA 5
 Db 12 EQQEA 16

Search completed: April 23, 2003, 13:43:49
 Job time : 4.75169 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:36:43 ; Search time 9.40449 Seconds
(without alignments)
183.999 Million cell updates/sec

Title: US-09-836-073-4
Perfect score: 102
Sequence: 1 QQQEAKICHOIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 4984

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	30.4	22	PH1359	Ig heavy chain DJ
2	30	29.4	25	A60286	heat-stable serine
3	29	28.4	18	PH1368	Ig heavy chain DJ
4	28	27.5	21	S78574	protein kinase C i
5	27	26.5	12	PH0771	T-cell receptor be
6	27	26.5	14	PH1626	Ig H chain V-D-J r
7	27	26.5	15	PH0789	T-cell receptor al
8	27	26.5	19	A05305	hemoglobin beta-2
9	27	26.5	21	S78575	protein kinase C i
10	27	26.5	22	PH1325	Ig heavy chain DJ
11	27	26.5	23	PH1725	Ig heavy chain V r
12	27	26.5	24	PH1696	Ig heavy chain V r
13	27	26.5	25	C57001	endo-1,4-beta-xyla
14	26	25.5	14	PH1598	Ig H chain V-D-J r
15	26	25.5	16	E53284	T-cell receptor be
16	26	25.5	18	PH1629	Ig H chain V-D-J r
17	26	25.5	19	B53145	high conductance c
18	26	25.5	22	I77373	gene N-ras protein
19	26	25.5	24	B53524	ubiquinol-cytochro
20	26	25.5	25	PH1700	Ig heavy chain V r
21	25	24.5	12	G64003	hypothetical prote
22	25	24.5	15	PH1366	Ig heavy chain DJ
23	25	24.5	17	A61211	anantin - Streptom
24	25	24.5	24	PT0258	Ig heavy chain CDR
25	25	24.5	25	PH1686	Ig heavy chain V r
26	24	23.5	12	S57570	T cell receptor V-
27	24	23.5	14	S57569	T cell receptor V-
28	24	23.5	14	S57638	T cell receptor V-
29	24	23.5	16	MTDFB5	melanotropin beta

30	24	23.5	18	1	MTDFB5	melanotropin beta
31	24	23.5	21	2	PH1730	Ig heavy chain V r
32	24	23.5	21	2	PH1730	proteinase inhibitor
33	24	23.5	22	2	PH1678	Ig heavy chain V r
34	24	23.5	22	2	PH1679	Ig heavy chain V r
35	24	23.5	23	2	PH1681	Ig heavy chain V r
36	24	23.5	23	2	PH1682	Ig heavy chain V r
37	24	23.5	23	2	PH1694	Ig heavy chain V r
38	24	23.5	23	2	PH1707	Ig heavy chain V r
39	24	23.5	23	2	PH1722	Ig heavy chain V r
40	24	23.5	23	2	PH1724	Ig heavy chain V r
41	24	23.5	23	2	PH1727	Ig heavy chain V r
42	24	23.5	23	2	PH1723	Ig heavy chain V r
43	24	23.5	24	2	PH1683	Ig heavy chain V r
44	24	23.5	24	2	PH1685	Ig heavy chain V r
45	24	23.5	24	2	PH1698	Ig heavy chain V r

ALIGNMENTS

RESULT 1
PH1359
Ig heavy chain DJ region (clone C178-121) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PH1359
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor
A:Reference number: PH1302; MUID:93094761; PMID:1460419
A:Accession: PH1359
A:Molecule type: DNA
A:Residues: 1-22 <WAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 30.4%; Score 31; DB 2; Length 22;
Best Local Similarity 44.4%; Pred. NO. 1.5e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 8 CHQIEYFG 16
I::||:|
DB 6 CYENYYFG 14

RESULT 2
A60286
heat-stable serine proteinase (EC 3.4.21.-) - Thermomonospora fusca (strain YX) (fr
N:Alternate names: YX-proteinase
C:Species: Thermomonospora fusca
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Dec-1996
C:Accession: A60286
R:Kirstjansson, M.M.; Kinsella, J.E.
Int. J. Pept. Protein Res. 36, 201-207, 1990
A:Title: Heat stable proteinase from Thermomonospora fusca. Characterization as a s
A:Reference number: A60286; MUID:91107200; PMID:2132918
A:Accession: A60286
A:Molecule type: protein
A:Residues: 1-25 <KRI>
C:Keywords: extracellular protein; hydrolase; serine proteinase

Query Match 29.4%; Score 30; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. NO. 2.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 13 YFGDF 18
||||:
DB 10 YFGNY 15

RESULT 3
PH1368

```

Db      6 QYEQYFG 12
| | | | |
RESULT 6
PH1626
Ig H chain V-D-J region (clone B-less 118) ~ mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C/Accession: PH1626
R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 176, 317-329, 1993
A/Title: Molecular characterization of transgene-induced immunodeficiency in B-less
A/Reference number: PH1580; MUID:93301609; PMID:8315387
A/Accession: PH1626
A/Molecule type: DNA
A/Residues: 1-14 <LEV>
A:Experimental source: bone marrow pre-B lymphocyte
C/Keywords: immunoglobulin

Query Match      26.5%; Score 27; DB 2; Length 14;
Best Local Similarity 40.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      8 CHQIEYFG 17
| : | | |
Db      1 CMRAXYVG 10

RESULT 7
PH0789
T-cell receptor alpha chain (E22 V-alpha-4.delta-7R) - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C/Accession: PH0789
R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A/Title: T cell receptor genes in a series of class I major histocompatibility comp
allelic exclusion and antigen-specific repertoire.
A/Reference number: PH0746; MUID:92078846; PMID:1836010
A/Accession: PH0789
A/Molecule type: mRNA
A/Residues: 1-15 <CAS>
A/Cross-references: EMBL:X60894
A:Experimental source: T lymphocyte
C/Keywords: T-cell receptor

Query Match      26.5%; Score 27; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      9 HQIEYFG 16
| : | | |
Db      8 HGLQYFG 15

RESULT 8
A03305
hemoglobin beta-2 chain - Indian spiny-tailed lizard (fragment)
C/Species: Uromastix hardwickii (Indian spiny-tailed lizard)
C/Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 01-Aug-1997
C/Accession: A03305
R:Naqvi, S.; Zaidi, Z.H.; von Bahr-Lindstrom, H.; Carlquist, M.; Jornvall, H.
FEBS Lett. 162, 290-295, 1983
A/Reference number: A91314; MUID:84029159; PMID:6628672
A/Accession: A03305
A/Molecule type: protein
A/Residues: 1-19 <NAQ>
C:Superfamily: globin; globin homology
C/Keywords: erythrocyte; oxygen carrier

Query Match      26.5%; Score 27; DB 2; Length 19;
Best Local Similarity 80.0%; Pred. No. 5.8e+02;

```

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 YFGDF 18
:||||
Db 1 PFGDF 5

RESULT 9

S78575
protein kinase C inhibitor KCIP-1 isoform eta - sheep (fragment)
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 21-Nov-1998
C:Accession: S78575
R:Martin, H.; Patel, Y.; Jones, D.; Howell, S.; Robinson, K.; Aitken, A.
FEBS Lett. 331, 296-303, 1993
A:Title: Antibodies against the major brain isoforms of 14-3-3 protein. An antibody specific for the major brain isoform of 14-3-3 protein. PMID:8375512
A:Reference number: S38299; MUID:93387487; PMID:8375512
A:Accession: S78575
A:Molecule type: protein
A:Residues: 1-21 <NAR>
A:Note: the acetylated initiator methionine is removed and the Gly undergoes a further modification
C:Superfamily: 14-3-3 protein
C:Keywords: acetylated amino end
F:/Modified site: acetylated amino end (Gly) #status experimental

Query Match 26.5%; Score 27; DB 2; Length 21;
Best Local Similarity 41.7%; Pred. No. 6.4e+02;
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 QEAKICHQIEY 14
|:|:|:|
Db 8 QRARLAQAEY 19

RESULT 10

PHI325
Ig heavy chain DJ region (clone C199-121) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PHI325
R:Wasserman, R.; Gallili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma. PMID:1460419
A:Reference number: PHI302; MUID:93094761; PMID:1460419
A:Accession: PHI325
A:Molecule type: DNA
A:Residues: 1-22 <WAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 26.5%; Score 27; DB 2; Length 22;
Best Local Similarity 66.7%; Pred. No. 6.7e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 IEYFYG 16
|:|:|
Db 9 IHVYFG 14

RESULT 11

PHI725
Ig heavy chain V region (clone GCC-5) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PHI725
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PHI1675; MUID:93301607; PMID:8315385
A:Accession: PHI725
A:Molecule type: mRNA
A:Residues: 1-23 <MCH>
A:Experimental source: B cell

A:Note: the authors translated the codon ACA for residue 13 as Ala
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 26.5%; Score 27; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 7e+02;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 10 QIEYFYGDF 18
|:|:|
Db 12 ETRYVYGSY 20

RESULT 12

PHI696
Ig heavy chain V region (clone NP-7-9) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PHI696
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PHI1675; MUID:93301607; PMID:8315385
A:Accession: PHI696
A:Molecule type: mRNA
A:Residues: 1-24 <MCH>
A:Experimental source: B cell
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 26.5%; Score 27; DB 2; Length 24;
Best Local Similarity 42.9%; Pred. No. 7.3e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 10 QIEYFYG 16
|:|:|
Db 12 EVAYYFG 18

RESULT 13

CS7001
endo-1,4-beta-xylanase (EC 3.2.1.8) 4 - Streptomyces roseiscleroticus (fragment)
C:Species: Streptomyces roseiscleroticus
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 23-May-1997
C:Accession: CS7001
R:Grabski, A.C.; Forrester, I.T.; Patel, R.; Jeffries, T.W.
Protein Expr. Purif. 4, 120-129, 1993
A:Title: Characterization and N-terminal amino acid sequences of beta-(1-4)endoxylase from Streptomyces roseiscleroticus. PMID:8471845
A:Reference number: A57001; MUID:93229899; PMID:8471845
A:Accession: CS7001
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-25 <GRA>
A:Experimental source: strain NRRL B-11019
A:Note: sequence extracted from NCBI backbone (NCBIP:130009)
C:Function:
A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
A:Pathway: xylan degradation
C:Superfamily: Streptomyces endo-1,4-beta-xylanase A; Streptomyces endo-1,4-beta-xylosidase; extracellular protein; glycosidase; hydrolase; polysaccharide degrading

Query Match 26.5%; Score 27; DB 2; Length 25;
Best Local Similarity 71.4%; Pred. No. 7.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYFYG 16
|:|:|
Db 11 QSGYFYG 17

RESULT 14

PHI598
Ig H chain V-D-J region (wild-type clone 306) - mouse (fragment)

C:Species: Mus musculus (house mouse)
 C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
 C:Accession: PH1598
 R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
 J. Exp. Med. 178, 317-329, 1993
 A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
 A:Reference number: PH1580; MUID:93301609; PMID:8315387
 A:Accession: PH1598
 A:Molecule type: DNA
 A:Residues: 1-14 <LEV>
 A:Experimental source: Bone marrow pre-B lymphocyte
 C:Keywords: immunoglobulin

Query Match 25.5%; Score 26; DB 2; Length 14;
 Best Local Similarity 50.0%; Pred. No. 6.3e+02;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Oy 8 CHQIEYFEGD 17
 : ||||
 Db 1 CAKDGYYGGD 10

RESULT 15
 E53284
 T-cell receptor beta 2 chain J region, Jbeta2.2 - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
 C:Accession: E53284
 R:Harindranath, N.; Alexander, C.B.; Mage, R.G.
 Mol. Immunol. 28, 881-888, 1991
 A:Title: Evolutionarily conserved organization and sequences of germline diversity and J
 A:Reference number: A53284; MUID:91342695; PMID:1678859
 A:Accession: E53284
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-16 <HAR>
 A:Cross-references: GB:S60737; NID:g233916; PIDN:AAB19521.1; PID:g233921
 A:Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBIP:60743)
 C:Keywords: T-cell receptor

Query Match 25.5%; Score 26; DB 2; Length 16;
 Best Local Similarity 66.7%; Pred. No. 7.2e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 12 EYFEGD 17
 : ||||
 Db 4 QLYFGD 9

Search completed: April 23, 2003, 13:48:54
 Job time : 10.4045 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: April 23, 2003, 13:43:09 ; Search time 9.91011 Seconds
(without alignments)
145.542 Million cell updates/sec

Title: US-09-836-073-4

Perfect score: 102

Sequence: 1 QQQEAKICHQIEYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 61792

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	102	100.0	18	9	US-09-836-073-4
2	90	88.2	18	9	US-09-836-073-5
3	87	85.3	17	9	US-09-836-073-13
4	87	85.3	18	9	US-09-836-073-1
5	87	85.3	18	9	US-09-836-073-14
6	81	79.4	18	9	US-09-836-073-9
7	79	77.5	18	9	US-09-836-073-11
8	79	77.5	18	9	US-09-836-073-12
9	78	76.5	18	9	US-09-836-073-10
10	76.5	75.0	19	9	US-09-836-073-16
11	75	73.5	18	9	US-09-836-073-2
12	75	73.5	18	9	US-09-836-073-3
13	72	70.6	18	9	US-09-836-073-7
14	72	70.6	18	9	US-09-836-073-15
15	71	69.6	18	9	US-09-836-073-8
16	62	60.8	16	9	US-09-836-073-19
17	56	54.9	18	9	US-09-836-073-6
18	44	43.1	18	9	US-09-836-073-17
19	38.5	37.7	18	9	US-09-836-073-18

20	30.5	29.9	23	9	US-09-813-153-291	Sequence 291, App
21	30	29.4	25	9	US-10-097-065-405	Sequence 405, App
22	29	28.4	14	8	US-08-424-550B-455	Sequence 455, App
23	29	28.4	17	10	US-09-864-761-47213	Sequence 47213, A
24	29	28.4	19	9	US-10-153-159-6	Sequence 6, Appl
25	29	28.4	19	9	US-10-153-159-51	Sequence 51, Appl
26	29	28.4	19	9	US-10-153-159-53	Sequence 53, Appl
27	29	28.4	19	9	US-10-153-159-54	Sequence 54, Appl
28	29	28.4	19	9	US-10-153-176-6	Sequence 6, Appl
29	29	28.4	19	9	US-10-153-176-51	Sequence 51, Appl
30	29	28.4	19	9	US-10-153-176-53	Sequence 53, Appl
31	29	28.4	19	9	US-10-153-176-54	Sequence 54, Appl
32	29	28.4	20	9	US-09-986-480-269	Sequence 269, App
33	28	27.5	15	10	US-09-073-009-51	Sequence 51, Appl
34	28	27.5	15	10	US-09-023-588-51	Sequence 51, Appl
35	28	27.5	15	10	US-09-793-306-51	Sequence 51, Appl
36	28	27.5	18	9	US-10-084-813-139	Sequence 139, App
37	28	27.5	18	9	US-10-084-813-140	Sequence 140, App
38	28	27.5	18	9	US-10-084-813-141	Sequence 141, App
39	28	27.5	21	9	US-09-880-748-3010	Sequence 3010, Ap
40	28	27.5	25	10	US-09-864-761-43809	Sequence 43809, A
41	27	26.5	15	9	US-09-989-919-123	Sequence 123, App
42	27	26.5	17	10	US-09-864-761-33337	Sequence 33337, A
43	27	26.5	18	10	US-09-840-009-13	Sequence 13, Appl
44	27	26.5	19	9	US-10-153-159-49	Sequence 49, Appl
45	27	26.5	19	9	US-10-153-159-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match 100.0% ; Score 102; DB 9; Length 18;
Best Local Similarity 100.0% ; Pred. No. 9.5e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQQEAKICHQIEYFGDF 18
Db 1 QQQEAKICHQIEYFGDF 18

RESULT 2
US-09-836-073-5
; Sequence 5, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073


```
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-5
```

```
Query Match      88.2%; Score 90; DB 9; Length 18;
Best Local Similarity 88.9%; Pred. No. 7.3e-08;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 QOQEKICHQIEYFGDF 18
    ||||| ||||| ||||| |||||
DB 1 QOQEKICHQIEYFGDF 18
```

```
RESULT 3
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13
```

```
Query Match      85.3%; Score 87; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 4 EAKICHQIEYFGDF 18
    ||||| ||||| ||||| |||||
DB 3 EAKICHQIEYFGDF 17
```

```
RESULT 4
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1
```

```
Query Match      85.3%; Score 87; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHQIEYFGDF 18
    ||||| ||||| ||||| |||||
DB 4 EAKICHQIEYFGDF 18
```

```
RESULT 5
US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14
```

```
Query Match      85.3%; Score 87; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 4 EAKICHQIEYFGDF 18
    ||||| ||||| ||||| |||||
DB 4 EAKICHQIEYFGDF 18
```

```
RESULT 6
US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9
```

```
Query Match      79.4%; Score 81; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 4 EAKICHQIEYFGD 17
    ||||| ||||| ||||| |||||
DB 4 EAKICHQIEYFGD 17
```

```
RESULT 7
US-09-836-073-11
```

```
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match      77.5%; Score 79; DB 9; Length 18;
Best Local Similarity 93.3%; Pred. No. 3.9e-06;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHQIEYFGDF 18
   ||||| ||||| |||||
DB 4 EAKICHQIEYFGDF 18

RESULT 8
US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match      77.5%; Score 79; DB 9; Length 18;
Best Local Similarity 93.3%; Pred. No. 3.9e-06;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHQIEYFGDF 18
   ||||| ||||| |||||
DB 4 EAKICHQIEYFGDF 18

RESULT 9
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match      73.5%; Score 75; DB 9; Length 18;
Best Local Similarity 86.7%; Pred. No. 1.7e-05;
```

```
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match      76.5%; Score 78; DB 9; Length 18;
Best Local Similarity 93.3%; Pred. No. 5.6e-06;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHQIEYFGDF 18
   ||||| ||||| |||||
DB 4 EAKICHQIEYFGDF 18

RESULT 10
US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match      75.0%; Score 76.5; DB 9; Length 19;
Best Local Similarity 93.8%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 4 EAKICHQIEYFGDF 19
   ||||| ||||| |||||
DB 4 EAKICHQIEYFGDF 19

RESULT 11
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match      73.5%; Score 75; DB 9; Length 18;
Best Local Similarity 86.7%; Pred. No. 1.7e-05;
```

Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
:|||||:|||||
Db 4 EAQICQOIEYFGDF 18

RESULT 12

US-09-836-073-3

; Sequence 3, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-3

Query Match 73.5%; Score 75; DB 9; Length 18;

Best Local Similarity 80.0%; Pred. No. 1.7e-05;

Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
:|||||:|||||
Db 4 QAKICHOIQYFGDF 18

RESULT 13

US-09-836-073-7

; Sequence 7, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-7

Query Match 70.6%; Score 72; DB 9; Length 18;

Best Local Similarity 92.9%; Pred. No. 4.9e-05;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17
:|||||:|||||
Db 4 EAKICHOIEYQGD 17

RESULT 14

US-09-836-073-15

; Sequence 15, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Xenopus
US-09-836-073-15

Query Match 70.6%; Score 72; DB 9; Length 18;

Best Local Similarity 80.0%; Pred. No. 4.9e-05;

Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
:|||||:|||||
Db 4 DKICQIEYFGDF 18

RESULT 15

US-09-836-073-8

; Sequence 8, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-8

Query Match 69.6%; Score 71; DB 9; Length 18;

Best Local Similarity 86.7%; Pred. No. 7e-05;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
:|||||:|||||
Db 4 EAKICHOIEQQFGDF 18

Search completed: April 23, 2003, 13:52:09

Job time : 10.9101 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:38:28 ; Search time 8.49438 Seconds
(without alignments)
62.349 Million cell updates/sec

Title: US-09-836-073-4

Perfect score: 102

Sequence: 1 QQEKAKICHIYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 147762

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	85.3	18	4 US-09-316-630-3	Sequence 3, Appl1
2	87	85.3	18	4 US-09-316-630-4	Sequence 4, Appl1
3	35	34.3	24	4 US-09-082-279B-1195	Sequence 1195, Ap
4	35	34.3	24	4 US-09-315-304B-1195	Sequence 1195, Ap
5	30	29.4	7	1 US-08-166-930-15	Sequence 15, Appl
6	30	29.4	7	2 US-08-727-045A-15	Sequence 15, Appl
7	30	29.4	7	4 US-09-408-172-15	Sequence 15, Appl
8	30	29.4	11	2 US-08-618-696-20	Sequence 20, Appl
9	30	29.4	11	2 US-08-211-312-9	Sequence 9, Appl1
10	30	29.4	11	3 US-09-033-753-20	Sequence 20, Appl
11	30	29.4	11	3 US-08-472-285-9	Sequence 9, Appl1
12	30	29.4	11	4 US-08-472-929-9	Sequence 9, Appl1
13	29	28.4	10	2 US-08-618-696-11	Sequence 11, Appl
14	29	28.4	10	3 US-09-033-753-11	Sequence 11, Appl
15	29	28.4	11	2 US-08-618-696-7	Sequence 7, Appl1
16	29	28.4	11	3 US-09-033-753-7	Sequence 7, Appl1
17	29	28.4	14	2 US-08-637-759B-113	Sequence 113, App
18	29	28.4	14	3 US-08-871-355A-113	Sequence 113, App
19	29	28.4	14	4 US-09-201-945-113	Sequence 113, App
20	29	28.4	14	4 US-09-298-924-55	Sequence 55, Appl
21	29	28.4	14	4 US-08-469-260A-455	Sequence 455, App
22	29	28.4	15	4 US-08-743-168B-29	Sequence 29, Appl
23	29	28.4	15	5 PCT-US96-10435-29	Sequence 29, Appl
24	29	28.4	25	1 US-08-238-163-6	Sequence 6, Appl1
25	28	27.5	10	2 US-08-618-696-2	Sequence 2, Appl1
26	28	27.5	10	2 US-08-618-696-10	Sequence 10, Appl
27	28	27.5	10	3 US-09-033-753-2	Sequence 2, Appl1

28	27.5	10	3	US-09-033-753-10	Sequence 10, Appl
29	27.5	11	2	US-08-618-696-1	Sequence 1, Appl1
30	27.5	11	2	US-08-618-696-6	Sequence 6, Appl1
31	27.5	11	2	US-08-618-696-16	Sequence 16, Appl
32	27.5	11	2	US-08-618-696-19	Sequence 19, Appl
33	27.5	11	3	US-09-033-753-1	Sequence 1, Appl1
34	27.5	11	3	US-09-033-753-6	Sequence 6, Appl1
35	27.5	11	3	US-09-033-753-16	Sequence 16, Appl
36	27.5	11	3	US-09-033-753-19	Sequence 19, Appl
37	27.5	19	4	US-09-441-502B-61	Sequence 61, Appl
38	27.5	20	4	US-08-505-250-8	Sequence 8, Appl1
39	27.5	20	4	US-08-503-250-8	Sequence 8, Appl1
40	27.5	21	2	US-08-825-349-1	Sequence 1, Appl1
41	27.5	22	6	5281520-43	Patent No. 5281520
42	27.5	24	4	US-09-082-279B-1199	Sequence 1199, Ap
43	27.5	24	4	US-07-963-329A-73	Sequence 73, Appl
44	27.5	24	4	US-09-315-304B-1199	Sequence 1199, Ap
45	27.5	24	5	PCT-US92-09443A-73	Sequence 73, Appl

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match 85.3%; Score 87; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHIQIEYFGDF 18
| | | | | | | | | | | | | | | | | |
Db 4 EAKICHIQIEYFGDF 18

RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

;; PRIOR FILING DATE: 1997-10-06
;; PRIOR APPLICATION NUMBER: 08/321,427
;; PRIOR FILING DATE: 1994-10-11
;; PRIOR APPLICATION NUMBER: 60/086,527
;; PRIOR FILING DATE: 1998-05-22
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
;; OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 85.3%; Score 87; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 EAKICHQIEYFGDF 18
| | | | | | | | | | | | | | | | | |
Db 4 EAKICHQIEYFGDF 18

RESULT 3
US-082-279B-1195
; Sequence 1195, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1195
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-1195

Query Match 34.3%; Score 35; DB 4; Length 24;
Best Local Similarity 46.2%; Pred. No. 22;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 1 QOQEAQKICHQIEY 13
| : | | | | : |
Db 2 QEWQKIMHREFDY 14

RESULT 4
US-09-315-304B-1195
; Sequence 1195, Application US/09315304B
; Patent No. 6348568
; GENERAL INFORMATION:
; APPLICANT: Barney, S.
; APPLICANT: Guthrie, K.
; APPLICANT: Merutka, G.
; APPLICANT: Anwer, M.
; APPLICANT: Lambert, D.
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
; FILE REFERENCE: 7872-052
; CURRENT APPLICATION NUMBER: US/09/315,304B

;; CURRENT FILING DATE: 1999-05-20
;; PRIOR APPLICATION NUMBER: 09/082,279
;; PRIOR FILING DATE: 1998-05-20
;; NUMBER OF SEQ ID NOS: 1667
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1195
;; LENGTH: 24
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Core polypeptide
US-09-315-304B-1195

Query Match 34.3%; Score 35; DB 4; Length 24;
Best Local Similarity 46.2%; Pred. No. 22;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 1 QOQEAQKICHQIEY 13
| : | | | | : |
Db 2 QEWQKIMHREFDY 14

RESULT 5
US-08-166-930-15
; Sequence 15, Application US/08166930
; Patent No. 5599678
; GENERAL INFORMATION:
; APPLICANT: Kraus, Michael
; APPLICANT: Stuber, Werner
; TITLE OF INVENTION: Synthetic Peptides, Antibodies Against
; TITLE OF INVENTION: Them and Their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/166,930
; FILING DATE: 15-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 42 42 736.3
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 02481.1351-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-166-930-15

Query Match 29.4%; Score 30; DB 1; Length 7;
Best Local Similarity 80.0%; Pred. No. 2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 8 CHQIE 12
| | | | |

Db 1 CHQVE 5

RESULT 6

US-08-727-045A-15
; Sequence 15, Application US/08727045A
; Patent No. 5981697
; GENERAL INFORMATION:
; APPLICANT: Kraus, Michael
; APPLICANT: Stuber, Werner
; TITLE OF INVENTION: Synthetic Peptides, Antibodies Against
; TITLE OF INVENTION: Them and Their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: DC

; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/727,045A
; FILING DATE: 08-OCT-1996
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 42 42 736.3
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 05552.1351-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-727-045A-15

Query Match 29.4%; Score 30; DB 2; Length 7;
Best Local Similarity 80.0%; Pred. No. 2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQIE 12

Db 1 CHQVE 5

RESULT 7

US-09-408-172-15
; Sequence 15, Application US/09408172
; Patent No. 6441141
; GENERAL INFORMATION:
; APPLICANT: Kraus, Michael
; APPLICANT: Stuber, Werner
; TITLE OF INVENTION: Synthetic Peptides, Antibodies Against
; TITLE OF INVENTION: Them and Their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: DC

; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/408,172
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/727,045
; FILING DATE: 08-OCT-1996
; APPLICATION NUMBER: DE P 42 42 736.3
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 05552.1351-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-408-172-15

Query Match 29.4%; Score 30; DB 4; Length 7;
Best Local Similarity 80.0%; Pred. No. 2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQIE 12

Db 1 CHQVE 5

RESULT 8

US-08-618-696-20
; Sequence 20, Application US/08618696
; Patent No. 5861475
; GENERAL INFORMATION:
; APPLICANT: COOPER, JR., J. ALLEN D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210

; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,696
; FILING DATE: 20-MAR-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/995,269
; FILING DATE: 12/21/92
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UOAB:002/PAR
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-618-696-20

Query Match 29.4%; Score 30; DB 2; Length 11;
Best Local Similarity 44.4%; Pred. No. 62;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 HQIEYFGD 17
Db 1 NEASYFFGD 9

RESULT 9

US-08-211-312-9
Sequence 9, Application US/08211312
Patent No. 5986051
GENERAL INFORMATION:
APPLICANT: LABIGNE, AGNES
APPLICANT: CUSSAC, VALERIE
APPLICANT: FERRERO, RICHARD
TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
FOR THE REGULATION AND MATURATION OF UREASE AND THEIR USE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/211.312
FILING DATE: 01-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 91 12198
FILING DATE: 03-OCT-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR92/00921
FILING DATE: 02-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5986051man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-075-0XPCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-211-312-9

Query Match 29.4%; Score 30; DB 2; Length 11;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 AKICHQI 11
Db 1 AKICYEI 7

RESULT 10

US-09-033-753-20
Sequence 20, Application US/09033753
Patent No. 6017883
GENERAL INFORMATION:
APPLICANT: COOPER, JI., J. ALLEN D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
INHIBITION OF PHAGOCYTES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,753.
FILING DATE:

CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/618,696
FILING DATE: 20-MAR-1996
APPLICATION NUMBER: 07/995,269
FILING DATE: 12/21/92
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: USAB:002/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-033-753-20

Query Match 29.4%; Score 30; DB 3; Length 11;
Best Local Similarity 44.4%; Pred. No. 62;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 HQIEYFGD 17
Db 1 NEASYFFGD 9

RESULT 11

US-08-472-285-9
Sequence 9, Application US/08472285
Patent No. 6027878
GENERAL INFORMATION:
APPLICANT: LABIGNE, AGNES
APPLICANT: CUSSAC, VALERIE
APPLICANT: FERRERO, RICHARD
TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
FOR THE REGULATION AND MATURATION OF UREASE AND THEIR USE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/472,285
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/211,312
FILING DATE: 01-JUL-1994
APPLICATION NUMBER: FR 91 12198
FILING DATE: 03-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR92/00921
FILING DATE: 02-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 6027878man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-075-0XPCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-472-285-9

Query Match 29.4%; Score 30; DB 3; Length 11;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AKICHOI 11

Db 1 AKICYEI 7

RESULT 12

US-08-472-929-9
Sequence 9, Application US/08472929
Patent No. 6271017

GENERAL INFORMATION:
APPLICANT: LABIGNE, AGNES
APPLICANT: CUSSAC, VALERIE
APPLICANT: FERRERO, RICHARD
TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
FOR THE REGULATION AND MATURATION OF UREASE AND THEIR USE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,929
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/211,312
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR92/00921
FILING DATE: 02-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 6271017man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-075-0XPCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-472-929-9

Query Match 29.4%; Score 30; DB 4; Length 11;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AKICHOI 11

Db 1 AKICYEI 7

RESULT 13

US-08-618-696-11
Sequence 11, Application US/08618696
Patent No. 5861475

GENERAL INFORMATION:
APPLICANT: COOPER, Jr., J. ALLEN D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
INHIBITION OF PHAGOCYTES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/618,696
FILING DATE: 20-MAR-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/995,269
FILING DATE: 12/21/92
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UOAB:002/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:

; LENGTH: 10 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-618-696-11

Query Match 28.4%; Score 29; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 82;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYFGD 17
; 1:1111
Db 1 EASYFFGD 8

RESULT 14

US-09-033-753-11
; Sequence 11, Application US/09033753
; Patent No. 6017883

; GENERAL INFORMATION:

; APPLICANT: COOPER, JR., J. ALLEN D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/033,753
; FILING DATE:
; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/618,696
; FILING DATE: 20-MAR-1996
; APPLICATION NUMBER: 07/995,269
; FILING DATE: 12/21/92

; ATTORNEY/AGENT INFORMATION:

; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UOAB:002/PAR

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 512-320-7200
; TELEFAX: 512-474-7577
; TELEX: NOT APPLICABLE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-09-033-753-11

Query Match 28.4%; Score 29; DB 3; Length 10;
Best Local Similarity 50.0%; Pred. No. 82;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYFGD 17
; 1:1111
Db 1 EASYFFGD 8

RESULT 15

US-08-618-696-7
; Sequence 7, Application US/08618696

; Patent No. 5861475
; GENERAL INFORMATION:
; APPLICANT: COOPER, JR., J. ALLEN D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210

; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,696
; FILING DATE: 20-MAR-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/995,269
; FILING DATE: 12/21/92

; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UOAB:002/PAR

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-320-7200
; TELEFAX: 512-474-7577
; TELEX: NOT APPLICABLE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-618-696-7

Query Match 28.4%; Score 29; DB 2; Length 11;
Best Local Similarity 50.0%; Pred. No. 91;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYFGD 17
; 1:1111
Db 2 EASYFFGD 9

Search completed: April 23, 2003, 13:50:25
Job time : 8.49438 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:28:18 ; Search time 27.2022 Seconds
(without alignments)
88.173 Million cell updates/sec

Title: US-09-836-073-4

Perfect score: 102
Sequence: 1 QQQEAKICHOIEYYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 320064

```
Minimum DB seq length: 0
Maximum DB seq length: 25
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

Database : A_Geneseq_101002.*

1:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SID52/qcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SID52/qcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SID52/qcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23:	/SID52/qcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	87	85.3	18	21	AAY52200	Human la autoantigen
2	73	71.6	21	14	AAR43394	Lay/SSB epitope 17.
3	37	36.3	10	23	ABB84046	Transglutaminase 1
4	36	35.3	23	22	AAU00924	VH ligand-binding
5	35	34.3	24	22	AAY98809	Core polypeptide f
6	35	34.3	24	22	ABB01217	Viral DPL178/107-11
7	35	34.3	24	22	ABB02668	Viral core polypep
8	35	34.3	24	22	AAU33763	DPL178-like/DPL107-1
9	35	34.3	24	22	AAB78210	Core polypeptide T
10	31	30.4	8	14	AAR43395	Lay/SSB epitope 24.

11	31	30.4	9	20	AAV10410	T cell epitope/MHC
12	31	30.4	10	23	AAE22202	Murine MC-1 antio
13	31	30.4	11	20	AAW95255	Anti-progesterone
14	31	30.4	16	22	AAW98909	Vaccine related MH
15	31	30.4	18	20	AAW93332	Human DPD protein
16	31	30.4	19	18	AAW27677	Coagulation inhibi
17	31	30.4	23	23	AAU99805	Biomembrane perme
18	30.5	29.9	23	20	AAV25901	Human secreted pro
19	30	29.4	7	15	AAR56273	Peptide antigen us
20	30	29.4	11	14	AAR34400	Fragment of Helico
21	30	29.4	11	15	AAR56307	Modified Influenza
22	30	29.4	15	18	AAW36481	VAT1 homologous pe
23	30	29.4	21	22	AAW48862	Mutant human insul
24	30	29.4	21	23	ABG66336	IgE Fcεpsilon RI b
25	30	29.4	22	20	AAV19621	SEQ ID NO 339 from
26	30	29.4	25	20	AAV36636	Fragment of human
27	29	28.4	9	23	AAU92267	PHOR1-F5D6 peptide
28	29	28.4	9	23	AAU92288	PHOR1-F5D6 peptide
29	29	28.4	9	23	AAU92398	PHOR1-F5D6 peptide
30	29	28.4	9	23	AAU92773	PHOR1-F5D6 peptide
31	29	28.4	9	23	AAU92875	PHOR1-F5D6 peptide
32	29	28.4	10	15	AAR56297	Synthetic modified
33	29	28.4	10	23	AAU92318	PHOR1-F5D6 peptide
34	29	28.4	10	23	AAU92332	PHOR1-F5D6 peptide
35	29	28.4	10	23	AAU92548	PHOR1-F5D6 peptide
36	29	28.4	10	23	AAU92639	PHOR1-F5D6 peptide
37	29	28.4	10	23	AAU92744	PHOR1-F5D6 peptide
38	29	28.4	10	23	AAU92821	PHOR1-F5D6 peptide
39	29	28.4	10	23	AAU92931	PHOR1-F5D6 peptide
40	29	28.4	10	23	AAU92942	PHOR1-F5D6 peptide
41	29	28.4	11	15	AAR56292	Synthetic derivati
42	29	28.4	12	21	AAV93796	Reactive peptide w
43	29	28.4	12	21	AAV93875	Reactive peptide w
44	29	28.4	14	17	AAR90665	S. acidocaldarius
45	29	28.4	14	21	AAR09328	Hepatitis GB virus

ALIGNMENTS

RESULT 1
AA52200
ID AAY52200 standard; peptide; 18 AA.

XX	AA52200;
XX	AC
XX	14-MAR-2000 (first entry)
DT	Human la autoantigen peptide (IAP).
XX	
DE	

La autoantigen; LAP; internal ribosome entry site; IRES; translation;
viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
coronavirus; hepatitis virus; rhadovirus; adenovirus; coxsackie virus;
parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
vesicular stomatitis virus.

XX
NS
Homo sapiens

XX PN WQ9961613-A2.

XX
PD 02-DEC-1999

XX 21-MAY-1999: 99WO-US11281.

XX
PR 22-MAY-1998: 98US-0086527.

XX PA (REGC) UNIV CALIFORNIA.

Das S. Dasgupta A:

XX
D8
WPT: 2000-062712/05.

XX XX

PT New yeast inhibitory peptide useful for inhibiting viral protein
 PT translation and replication -
 XX
 PS Claim 5; Page 57; 81pp; English.
 CC This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (IRNA see A4245200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rhadovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.
 XX

Sequence 18 AA;
 Query Match 85.3%; Score 87; DB 21; Length 18;
 Best Local Similarity 100.0%; Pred. No. 2e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 EAKICHQIEYYFGDF 18
 |||||
 DB 4 EAKICHQIEYYFGDF 18

RESULT 2

AA43394
 ID AA43394 standard; peptide; 21 AA.

AC AA43394;

DT 12-MAY-1994 (first entry)

DE La/SSB epitope 17.

Linear epitope: 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

OS Homo sapiens.

PN WO9321223-A.

PD 28-OCT-1993.

PF 13-APR-1993; 93WO-US03484.

PR 13-APR-1992; 92US-0867819.

PA (OKLA) UNIV OKLAHOMA STATE.

PI Harley JB;

DR WPI; 1993-351658/44.

XX New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus

PS Claim 1; Page 30; 43pp; English.

CC The sequences given in AA43391-562 are linear epitopes which are

CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B',
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polynucleotide
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.
 XX

Sequence 21 AA;

Query Match 71.6%; Score 73; DB 14; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 ICHQIEYYFGDF 18
 |||||

DB 1 ICHQIEYYFGDF 12

RESULT 3

ABB84046

ID ABB84046 standard; peptide; 10 AA.

AC ABB84046;

DT 21-AUG-2002 (first entry)

DE Transglutaminase inhibitory peptide cr type #16.

XX Transglutaminase inhibitor; Factor XIII inhibitor; XIIIa inhibitor;
 KW ophthalmological; antiinflammatory; antirheumatic; antiarthritic;
 KW thrombolytic; neuroprotective; nootropic; antiseborrheic; dermatological;
 KW cytostatic; anti-HIV; antipsoriatic; cataract; inflammatory disease;
 KW arthritis; thrombosis; Alzheimer's disease; Huntington's chorea;
 KW acne; cancer; HIV infection; psoriasis.

XX Unidentified.

OS

PN WO200236798-A2.

PD 10-MAY-2002.

PF 02-NOV-2001; 2001WO-EP12727.

PR 03-NOV-2000; 2000DE-1054687.

PA (NZYM-) N ZYME BIOTEC GMBH.

PI Fuchsbauser H, Pasternack R, Zotzel J;

DR WPI; 2002-444364/47.

XX New amino acid or peptide derivatives or analogs, are selective
 PT transglutaminase inhibitors useful e.g. for treating cataract,
 PT inflammatory diseases, rheumatoid arthritis, thrombosis, Alzheimer's
 PT disease and cancer

PS Disclosure; Page 13; 44pp; German.

XX This invention describes novel amino acid or peptide derivatives or
 CC analogues (I), containing a modified side-chain (e.g. containing a formyl

CC group) which are transglutaminase inhibitors and Factor XIII/XIIIa
 CC inhibitors. The products of the invention have ophthalmological,
 CC antiinflammatory, antirheumatic, antiarthritic, thrombolytic,
 CC neuroprotective, nootropic, antiseborrheic, dermatological, cytostatic,
 CC anti-HIV and antipsoriatic activity. (1) are transglutaminase inhibitors,
 CC especially inhibitors of crosslinking of proteins or peptides
 CC (specifically fibrin and/or alpha₂-plasmin inhibitor), incorporation of
 CC primary amines in proteins and peptides, hydrolysis of the
 CC gamma-carboxamido group of glutamine residues bound in proteins or
 CC peptides, blood factor XIII/XIIIa and mammalian, human, tissue, liver,
 CC brain, eye lens, keratinocyte, epidermal, prostate, plant, parasitic
 CC and/or bacterial transglutaminases. The products of the invention can be
 CC used for treating cataract, inflammatory diseases, rheumatoid arthritis,
 CC chronic arthritis, thrombosis, Alzheimer's disease, Huntington's chorea,
 CC acne, cancer (by induction of apoptosis), HIV infections and psoriasis.
 CC (1) Are targeted and specific transglutaminase inhibitors, which can
 CC inhibit a specific type of transglutaminase in the human or animal body
 CC without affecting other transglutaminases. ABB84001-ABB84049 represent
 CC transglutaminase inhibitors described in the method of the invention.
 XX
 SQ Sequence 10 AA;

Query Match 36.3%; Score 37; DB 23; Length 10;
 Best Local Similarity 66.7%; Pred. No. 16;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 QOEAKICHQ 10
 Db 1 QKQAPICHQ 9

RESULT 4
 AAU00924
 ID AAU00924 standard; Peptide; 23 AA.

AC AAU00924;

DT 18-MAY-2001 (first entry)

DE VH ligand-binding domain variant CDR3 region #31.

XX Human; monoclonal antibody A6; VH ligand-binding domain; variable region;
 KW parental ligand binding molecule; PLBM; framework region; FR; CDR;
 KW combinatorial library; diagnosis; therapy; surface protein; pesticide;
 KW cell-receptor; carbohydrate; lipid; antibiotic; steroid; hormone; dye;
 KW herbicide; polymer; immunoglobulin; CDR3 region; M2 immunoglobulin;
 KW panning.

XX Homo sapiens.

XX WO200118058-A2.

PN 15-MAR-2001.

PF 07-SEP-2000; 2000WO-CA01027.

PR 07-SEP-1999; 99CA-2282179.

PR 04-NOV-1999; 99US-0163546.

PA (NOVO-) NOVOPHARM BIOTECH INC.

PI Kaplan H, Entwistle J, Tanha J, Narang S, Dan M;

DR WPI; 2001-235191/24.

XX Combinatorial libraries including phage display library comprises
 PT variants of immunoglobulin VH fragments which comprises the framework
 PT regions of wild-type or modified immunoglobulin VH domain of human A6
 PT antibody.

XX Disclosure; Page 25; 133pp; English.

XX The sequence represents a variant CDR3 region of a parental VH

CC ligand-binding domain of human monoclonal antibody A6, isolated by
 CC panning against M2 immunoglobulin. Variants of a parental ligand
 CC binding molecule (PLBM) comprise an immunoglobulin (Ig) VH fragment
 CC including the framework (FR) regions of the wild-type/modified Ig VH
 CC domain of human monoclonal antibody A6. The variant differs from PLBM at
 CC residues constituting part of one of the CDRs of PLBM. A library
 CC expressing binding domains can be created by cloning a parental DNA
 CC sequence that encodes a parental domain, in order to produce parental
 CC clones, and replacing a variable region of the clones with a variant DNA
 CC sequence. A library of genetic packages can then be generated, each
 CC package having a surface and a surface protein binding domain expressed
 CC by the variant DNA. Combinatorial libraries are useful for generating
 CC forms of recombinant antibody fragments that specifically bind target
 CC ligands including natural and synthetic molecules and macromolecules to
 CC be used in diagnostic and/or therapeutic processes. Examples include
 CC enzymes, cell-receptors, carbohydrates, lipids, and organic-based and
 CC inorganic-based molecules such as antibiotics, steroids, hormones,
 CC pesticides, herbicides, dyes and polymers.

XX Sequence 23 AA;

Query Match 35.3%; Score 36; DB 22; Length 23;
 Best Local Similarity 45.5%; Pred. No. 58;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 CHQIEYFGDF 18
 Db 11 CFWDYKFSDF 21

RESULT 5
 AAY89809
 ID AAY89809 standard; peptide; 24 AA.

AC AAY89809;

DT 23-MAY-2000 (first entry)

DE Core polypeptide fragment T No. 1377.

XX Retrovirus; hybrid polypeptide; enhancer; gp41; envelope protein; HIV-1;
 KW HIV-2; SIV; pharmacokinetic; half-life; growth factor; cytokine; viral;
 KW anti-fusogenic; differentiation factor; interleukin; interferon;
 KW colony stimulating factor; hormone; angiogenic factor.

OS Unidentified.

XX WO9959615-A1.

XX 25-NOV-1999.

PF 20-MAY-1999; 99WO-US11219.

PR 20-MAY-1998; 98US-0082279.

PA (TRIM-) TRIMERIS INC.

PI Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;

DR WPI; 2000-136792/12.

XX A new hybrid polypeptide with enhanced pharmacokinetic properties
 PT comprises enhancer sequence -

XX Disclosure; Page 44; 124pp; English.

XX The invention relates to hybrid polypeptides comprising enhancer peptide
 CC sequence linked to core polypeptides. The enhancer polypeptides are
 CC derived from various retroviral envelope (gp41) protein sequences,
 CC especially from HIV-1, HIV-2 and SIV. The enhancer peptides enhance the
 CC pharmacokinetic properties such as increasing the half-life of any core
 CC polypeptide that they are linked to. The core polypeptides are any
 CC polypeptide that may be introduced into a living system and that can

CC function as a pharmacologically useful peptide for the treatment or
 CC prevention of a disease. The core polypeptides are bioactive peptides
 CC selected from a growth factor, cytokine, differentiation factor,
 CC interleukin, interferon, colony stimulating factor, hormone or
 CC angiogenic factor. The peptides of the invention can be used for
 CC inhibiting viral infection and can be used in anti-viral and
 CC anti-fusogenic treatments. Sequences AAY8651-Y90055 represent core
 CC polypeptide fragments that can be used in the invention. Some sequences
 CC among those indicated also comprise enhancer fragments at terminal ends
 CC and form hybrid polypeptides.

xx SQ Sequence 24 AA;

Query Match 34.3%; Score 35; DB 21; Length 24;
 Best Local Similarity 46.2%; Pred. No. 88;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 1 QOQEAKEKICHOIEY 13
 I: | | | | | : |
 Db 2 QEWQKIMHREFDY 14

RESULT 6

ABBO1217
 ID ABB01217 standard; Peptide; 24 AA.

xx AC ABB01217;

xx DT 03-JAN-2002 (first entry)

xx DE Viral DP178/107-like region peptide T1377.

xx KW Human immunodeficiency virus; HIV; respiratory syncytial virus; RSV;
 KW virucide; heptad repeat region; transmembrane protein; gp41; HR1; HR2;
 infection.

xx OS Virididae.

xx FH Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal is substituted by Ac"

FT Modified-site 24

FT /note= "C-terminal amide"

xx PN WO200164013-A2.

xx PD 07-SEP-2001.

xx PF 07-FEB-2001; 2001WO-US03988.

xx PR 29-FEB-2000; 2000US-0515965.

xx PA (TRIM-) TRIMERIS INC.

xx PI Antczak JB, Delmedico MK, Erickson JB, Lambert DM, Sista P;

xx DR WPI; 2001-514829/56.

xx HT Heptad repeat region peptide analogs useful for inhibiting virus/cells
 PT fusion, useful for treating HIV and Respiratory Syncytial Virus
 PT infection -

xx PS Disclosure; Page 57; 587pp; English.

xx The invention relates to isolated analogues of the heptad repeat region
 CC peptides DP178 and DP107. DP178 and DP107 correspond to amino acids
 CC 638-673 (heptad repeat region HR2) and 558-595 (heptad repeat region
 CC HR1) respectively, of HIV-1LAI transmembrane protein gp41. The HR1
 CC and HR2 regions of proteins interact non-covalently with each other
 CC and/or with peptides derived from them. This interaction is required for
 CC normal infectivity of viruses such as RSV and HIV. The heptad
 CC repeat region peptide analogues may be used to inhibit respiratory
 CC syncytial virus (RSV) infection in a cell. They may also be used to

CC inhibit HIV infection. The present sequence is a peptide provided in
 CC the specification.

xx SQ Sequence 24 AA;

Query Match 34.3%; Score 35; DB 22; Length 24;
 Best Local Similarity 46.2%; Pred. No. 88;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 1 QOQEAKEKICHOIEY 13
 I: | | | | | : |
 Db 2 QEWQKIMHREFDY 14

RESULT 7

ABBO2668
 ID ABB02668 standard; Peptide; 24 AA.

xx AC ABB02668;

xx DT 03-JAN-2002 (first entry)

xx DE Viral core polypeptide, SEQ ID NO: 1195.

xx KW Human immunodeficiency virus; HIV; respiratory syncytial virus; RSV;
 KW virucide; heptad repeat region; transmembrane protein; gp41; HR1; HR2;
 infection.

xx OS Virididae.

xx PN WO200164013-A2.

xx PD 07-SEP-2001.

xx PF 07-FEB-2001; 2001WO-US03988.

xx PR 29-FEB-2000; 2000US-0515965.

xx PA (TRIM-) TRIMERIS INC.

xx PI Antczak JB, Delmedico MK, Erickson JB, Lambert DM, Sista P;

xx DR WPI; 2001-514829/56.

xx HT Heptad repeat region peptide analogs useful for inhibiting virus/cells
 PT fusion, useful for treating HIV and Respiratory Syncytial Virus
 PT infection -

xx PS Disclosure; Page 483; 587pp; English.

xx The invention relates to isolated analogues of the heptad repeat region
 CC peptides DP178 and DP107. DP178 and DP107 correspond to amino acids
 CC 638-673 (heptad repeat region HR2) and 558-595 (heptad repeat region
 CC HR1) respectively, of HIV-1LAI transmembrane protein gp41. The HR1
 CC and HR2 regions of proteins interact non-covalently with each other
 CC and/or with peptides derived from them. This interaction is required for
 CC normal infectivity of viruses such as RSV and HIV. The heptad
 CC repeat region peptide analogues may be used to inhibit respiratory
 CC syncytial virus (RSV) infection in a cell. They may also be used to
 CC inhibit HIV infection. The present sequence is a peptide provided in
 CC the specification.

xx SQ Sequence 24 AA;

Query Match 34.3%; Score 35; DB 22; Length 24;
 Best Local Similarity 46.2%; Pred. No. 88;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 1 QOQEAKEKICHOIEY 13
 I: | | | | | : |
 Db 2 QEWQKIMHREFDY 14

RESULT 8

AAU13763
ID AAU13763 standard; Peptide; 24 AA.

XX AC AAU13763;
XX DT 21-NOV-2001 (first entry)
XX DE DP178-like/DP107-like peptide T-1377.
XX KW Anti-retroviral; DP178-like; transmembrane protein gp41;
XX KW antifusogenic; antiviral; HIV transmission; mutant; mutelin.
XX OS Human immunodeficiency virus 1 isolate LAI.
XX OS Synthetic.

XX Key Location/Qualifiers
FH Modified-site 1
FT FT Modified-site 24 /note= "N-terminal is substituted by Ac"
FT FT Modified-site 24 /note= "C-terminal amide"
FT FT

XX WO200151673-A2.
XX 19-JUL-2001.
XX

XX 05-JUL-2000; 2000WO-US35727.
XX

XX 09-JUL-1999; 99US-0350841.
XX

XX (TRIM-) TRIMERIS INC.
XX

XX Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;
XX

XX WPI; 2001-442157/47.
XX

XX Identifying a compound that inhibits the formation of or disrupts a
PT DP107/DP178 complex, especially compounds with antifusogenic, antiviral
PT or intracellular modulatory activity, by detecting the formation of a
PT DP107/DP178 complex -
XX

XX Disclosure; Page 77; 259pp; English.
XX

XX The present invention relates to peptides which exhibit anti-retroviral
CC activity. The peptides of the invention (AAU12559-AAU14009) comprise
CC DP178-like and DP107-like peptides. The DP178 peptide corresponds
CC to amino acids 639-673 of the transmembrane protein gp41 from human
CC immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide
CC corresponds to amino acids 558-595 of gp41 from HIV-1LAI. The invention
CC also relates to a method of identifying compounds that inhibit the
CC formation of or disrupts a DP107/DP178 complex. The method comprises
CC detecting the formation of a DP107/DP178 complex, both in the presence
CC or absence of a test compound, in a reaction mixture containing DP107
CC and DP178 peptides. The method is useful for identifying compounds,
CC including small molecule compounds, which may themselves exhibit
CC antifusogenic, antiviral or intracellular modulatory activity. The
CC DP178-like/DP107-like peptides are useful to inhibit human and non-human
CC retroviral, particularly HIV, transmission to uninfected cells. The
CC present sequence represents one of the DP178-like/DP107-like peptides
CC of the invention.
XX

XX Sequence 24 AA;

Query Match 34.3%; Score 35; DB 22; Length 24;
Best Local Similarity 46.2%; Pred. No. 88;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 OQOEAKICHQIEY 13
I: | | | | | : |
Db 2 QEWQKIMHRFDY 14

RESULT 9

AAU43395
ID AAR43395 standard; peptide; 8 AA.

XX AC AAR43395;

AAU78210
ID AAB78210 standard; Peptide; 24 AA.

XX AC AAB78210;
XX DT 19-APR-2001 (first entry)
XX DE Core polypeptide T1377.
XX KW Core polypeptide; enhancer; antiviral; anti-HIV;
XX KW virucide; hepatotropic; antiinflammatory; hybrid polypeptide;
XX KW coiled-coil peptide interaction; fusion-related disorder;
XX KW bacterial infection; viral infection.
XX OS Unidentified.

XX WO200103723-A1.
XX

XX 18-JAN-2001.
XX

XX 10-JUL-2000; 2000WO-US18772.
XX

XX 09-JUL-1999; 99US-0350641.
XX

XX (TRIM-) TRIMERIS INC.
XX

XX Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;
XX

XX WPI; 2001-147136/15.
XX

XX New hybrid polypeptide, useful for preventing, treating and diagnosing
PT e.g. viral infections, comprises an enhancer peptide linked to a core
PT polypeptide -
XX

XX Disclosure; Page 58; 151pp; English.
XX

XX The present sequence is a core polypeptide which may be linked to
CC an enhancer peptide to form a novel hybrid polypeptide. The hybrid
CC polypeptide exhibits enhanced pharmacokinetic properties relative to
CC those exhibited by the core polypeptide when introduced into a living
CC system. It is used to increase the in vitro or ex vivo half-life of
CC the core polypeptide. The hybrid and core polypeptides can be used for
CC modulating fusogenic events and intracellular processes involving
CC coiled-coil peptide interactions. Other uses include preventing,
CC treating and/or diagnosing disorders involving fusion events (e.g.
CC modulation of neurotransmitter exchange and sperm-egg fusion),
CC intracellular processes involving coiled-coil peptides (e.g. bacterial
CC infections) and viral infections that involve cell-cell and/or
CC virus-cell fusion (e.g. viral infections caused by human
CC immunodeficiency virus, respiratory syncytial virus, Epstein-Barr
CC virus, hepatitis B virus, Mason-Pfizer virus and polio virus).
CC The enhancer peptide sequence increases the half-life and reduces the
CC clearance rate of therapeutic peptides, which increases their efficacy
CC and minimises the incidence and severity of adverse side effects.
CC In addition, this increases the sensitivity of the diagnostic procedure
CC in which they are used.
XX

XX Sequence 24 AA;

Query Match 34.3%; Score 35; DB 22; Length 24;
Best Local Similarity 46.2%; Pred. No. 88;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 OQOEAKICHQIEY 13
I: | | | | | : |
Db 2 QEWQKIMHRFDY 14

RESULT 10

AAR43395
ID AAR43395 standard; peptide; 8 AA.

XX AC AAR43395;

XX 12-MAY-1994 (first entry)
 DT La/SSB epitope 24.
 DE
 XX Linear; epitope: 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; rRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
 OS Homo sapiens.
 XX
 XX WO9321223-A.
 PN
 XX 28-OCT-1993.
 PD
 XX 13-APR-1993; 93WO-US03484.
 PR
 XX 13-APR-1992; 92US-0867819.
 PR
 XX (OKLA) UNIV OKLAHOMA STATE.
 PA
 XX Harley JB;
 PI
 XX WPI; 1993-351658/44.
 DR
 XX New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus
 XX
 PS Claim 1; Page 30; 43pp; English.
 XX
 CC The sequences given in AAR43391-562 are linear epitopes which are
 CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (rRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polynucleotide
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.
 XX
 SQ Sequence 8 AA;
 Query Match 30.4%; Score 31; DB 14; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 14 YFGDF 18
 |||||
 Db 1 YFGDF 5
 RESULT 11
 AAY10410
 ID AAY10410 standard; Peptide; 9 AA.
 XX
 AC AAY10410;
 XX
 DT 12-MAY-1999 (first entry)
 XX

DE T cell epitope/MHC ligand SEQ ID NO:340.
 XX
 KW Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;
 KW immunisation; tumour; infectious disease; immunotherapy; cancer;
 KW malignant melanoma; viral disease; hepatitis; AIDS.
 XX
 OS Synthetic.
 OS Simian virus.
 XX
 PN WO9902183-A2.
 XX
 XX 21-JAN-1999.
 PD
 XX 10-JUL-1998; 98WO-US14289.
 PF
 XX 10-DEC-1997; 97US-0988320.
 PR
 XX 10-JUL-1997; 97CA-2209815.
 PR
 XX (CTLI-) CTL IMMUNOTHERAPIES CORP.
 PA
 XX Kuendig TM, Simard JUL;
 PI
 XX WPI; 1999-120514/10.
 DR
 XX Inducing a cytotoxic T lymphocyte response - by maintaining a level
 PT of antigen in the lymphatic system of a mammal so as to provide a
 PT sustained CTL response, used to treat, e.g. AIDS
 XX
 PS Disclosure; Page 37; 199pp; English.
 XX
 CC The present invention describes a method of inducing and/or sustaining
 CC an immunological cytotoxic T lymphocyte (CTL) response in a mammal. The
 CC method comprises: (a) delivering an antigen to the mammal at a level to
 CC induce an immunological CTL response in the mammal; and (b) maintaining
 CC the level of the antigen in the mammal's lymphatic system to maintain
 CC the immunologic CTL response. The method can be used for the delivery of
 CC e.g. a differentiation antigen, a tumour-specific multilineage antigen,
 CC an embryonic antigen, an oncogene antigen, a mutated tumour-suppressor
 CC gene antigen, or a viral antigen. They can be used for the treatment of
 CC disease such as cancer, e.g. malignant melanoma or infectious disease,
 CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery
 CC to the lymphatic system provides for potent CTL stimulation that takes
 CC place in the milieu of the lymphoid organ, and it sustains stimulation
 CC that is necessary to keep CTL active, cytotoxic and recirculating
 CC through the body. AAY10071 to AAY10639 represent examples of peptide
 CC antigens given in the present invention.
 XX
 SQ Sequence 9 AA;
 Query Match 30.4%; Score 31; DB 20; Length 9;
 Best Local Similarity 75.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 4 EAKICHQI 11
 |||||
 Db 2 EAEIAHQI 9
 RESULT 12
 AAE22202
 ID AAE22202 standard; peptide; 10 AA.
 XX
 AC AAE22202;
 XX
 XX 25-JUL-2002 (first entry)
 DT
 XX Murine MC-1 antibody heavy chain variable domain (VH) CDR3 peptide.
 DE
 XX Chemokine construct; human immunodeficiency virus 1; allergic disease;
 KW skin disease; immunological disorder; autoimmune disease; psoriasis;
 KW multiple sclerosis; rheumatoid arthritis; inflammatory bowel disease;
 KW diabetes; skin inflammation; atopic dermatitis; inflammatory disease;
 KW inflammatory renal disease; HIV-1; transplant rejection; murine; CDR3;

KW heavy chain variable domain; VH; complementarity determining region 3;
 KW MC-1; antibody.
 XX Mus sp.
 OS
 XX WO200220615-A2.
 PN
 XX 14-MAR-2002.
 PD
 XX 10-SEP-2001; 2001WO-BP10433.
 PF
 XX 08-SEP-2000; 2000EP-0119694.
 PR
 XX 05-SEP-2001; 2001US-0948004.
 XX (MICR-) MICROMET AG.
 PA
 XX Mack M, Schloendorff D, Spring M;
 PI
 XX WPI; 2002-362240/39.
 DR
 XX Use of an antibody and/or chemokine construct that binds to a chemokine
 PT receptor, for eliminating cells latently infected with primate
 PT immunodeficiency virus, or treating, preventing and alleviating immune
 PT disorders -
 XX Disclosure; Page 117; 117pp; English.
 PS
 XX The invention relates to the use of an antibody and/or chemokine
 CC construct that binds a chemokine receptor for preparing a pharmaceutical
 CC composition for eliminating cells latently infected with a primate
 CC immunodeficiency virus (e.g. human immunodeficiency virus (HIV-1). They
 CC are used in gene therapy and as vaccines. The antibody and/or chemokine
 CC construct is also used for preparing a pharmaceutical composition for
 CC treating, preventing and/or alleviating immunological disorders
 CC including autoimmune diseases (e.g. multiple sclerosis, type I diabetes
 CC and rheumatoid arthritis), allergic diseases, skin diseases (e.g. skin
 CC inflammation, atopic dermatitis and psoriasis), inflammatory diseases
 CC such as inflammatory joint diseases (chronic arthritis), inflammatory
 CC renal diseases and inflammatory bowel diseases and graft versus host
 CC and transplant reactions. The present sequence is murine MC-1 antibody
 CC heavy chain variable domain (VH) complementarity determining region 3
 CC (CDR3) peptide.
 XX
 SQ Sequence 10 AA;
 Query Match 30.4%; Score 31; DB 23; Length 10;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 12 EYVGFDF 18
 Db III:I
 2 EYVGFDF 8
 RESULT 13
 AAW95255
 ID AAW95255 standard; peptide; 11 AA.
 AC AAW95255;
 XX
 XX 11-MAR-1999 (first entry)
 DT
 XX Anti-progesterone antibody Vh region (clone 1578/p2) CDR H3 sequence.
 DE
 XX ARM; protein selection; display; cell free system; genetic information;
 KW reverse transcription; single-chain antibody; RT-PCR; primer; drug;
 KW antibody-ribosome-mRNA complex; antibody engineering; progesterone;
 KW testosterone; human
 XX Homo sapiens.
 OS
 XX WO9854312-A1.
 PN
 XX

PD 03-DEC-1998.
 XX
 PF 28-MAY-1998; 98WO-GB01564.
 XX
 XX 28-FEB-1998; 98GB-0004195.
 PR
 XX 28-MAY-1997; 97GB-0010829.
 PR
 XX 26-NOV-1997; 97GB-0024850.
 XX
 PA (BABR-) BABRAHAM INST.
 XX
 XX He M, Taussig MJ;
 PI
 XX WPI; 1999-059832/05.
 DR
 XX In vitro display and evolution of proteins - by transcription and
 PT translation of DNA in a cell free system, selection and recovery of
 PT complexes and RT-PCR on RNA bound to ribosomes
 PT
 XX Example 13; Fig 19; 62pp; English.
 PS
 XX The invention relates to methods for the display and selection of
 CC proteins or peptides and for recovery of the genetic material encoding
 CC them. One method comprises (a) transcription and translation of DNA in a
 CC cell free system such that complexed particles are formed, each
 CC comprising at least one individual nascent protein or peptide or other
 CC DNA expression product associated with one or more ribosomes and the
 CC specific mRNA encoding the protein or peptide; (b) contacting the
 CC complexed particles with a ligand, antigen, antibody or other agent in
 CC order to select particles through binding to the protein or peptide
 CC product; and (c) recovering the genetic information encoding the protein
 CC or peptide as DNA by RT-PCR carried out on the mRNA while the latter
 CC remains bound to the complexed particle. The steps of display, selection
 CC and recovery can be repeated in consecutive cycles. The method is
 CC exemplified using single-chain antibody constructs as antibody-ribosome-
 CC mRNA complexes (ARMs). Methods in which the DNA is produced by RT-PCR,
 CC methods for making antibodies of human, mouse or rat are also provided.
 CC The methods can be used for the display and selection of single chain
 CC antibody fragments from libraries, antibody engineering, selection of
 CC human antibodies and selection of proteins from mRNA libraries. They can
 CC also be used to select ligands for combining sites or receptors, such
 CC ligands having potential uses as drugs or therapies. By carrying out the
 CC RT-PCR recovery step directly on the intact ribosome complex without
 CC prior dissociation to release the mRNA maximal efficiency and
 CC sensitivity can be obtained. Peptides AAW95247 to AAW95271 represent
 CC sequences of human anti-progesterone and anti-testosterone antibodies
 CC isolated from an immunised transgenic mouse by ARM display.
 XX
 SQ Sequence 11 AA;
 Query Match 30.4%; Score 31; DB 20; Length 11;
 Best Local Similarity 50.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 9 HQTEYFEG 16.
 Db :::::II
 3 YEIDWYFG 10
 RESULT 14
 AAW98909
 ID AAW98909 standard; Peptide; 16 AA.
 AC AAW98909;
 XX
 XX 07-DEC-2001 (first entry)
 DT
 XX Vaccine related MHC ligand peptide SEQ ID NO:12.
 DE
 XX Glutamic acid; glutamine; vaccine; major histocompatibility complex;
 KW MHC; immunomodulator; anti-allergic; endocrine; neuroprotectant;
 KW virucidal; bactericidal; antiparasitic; fungicidal; cytostatic;
 KW medicine; pharmaceutical; immune disorder; immune deficiency;
 KW autoimmune; hypersensitivity; allergy; graft rejection; infection;
 KW

KW hormonal disorder; central nervous system disease; cancer; melanoma;
 KW anti-melanoma vaccine; human immunodeficiency virus.

OS Rabies virus.

PN WO200170772-A2.

XX 27-SEP-2001.

XX 22-MAR-2001; 2001WO-FR00872.

XX 23-MAR-2000; 2000FR-0003711.

XX (FABR) FABRE MEDICAMENT SA PIERRE.

XX Klinguer-Hamour C, Corvaia N, Beck A, Goetsch L;

XX WPI; 2001-611470/70.

XX Stabilized pharmaceutical containing N-terminal glutamic acid or
 PT glutamine, useful e.g. in anti-melanoma vaccines, is an addition salt
 PT with strong acid

XX Claim 9; Page 31; 149pp; French.

XX The present invention describes a pharmaceutical compound (I) that
 CC contains an N-terminal glutamic acid (Glu) or glutamine (Gln) residue
 CC in the form of an addition salt with a strong, physiologically
 CC acceptable acid (II). Also described are: (a) a pharmaceutical
 CC composition containing at least one (I); (b) a vaccine containing at
 CC least one (I) where this is a major histocompatibility complex (MHC)
 CC ligand (Ia); (c) a method for in vitro diagnosis of diseases associated
 CC with the presence of (Ia); (d) a kit for method (c) that includes a (Ia);
 CC and (e) a process for preparing (I). (I) has immunomodulator, endocrine,
 CC antiallergic, neuroprotectant, virucidal, bactericidal, antiparasitic,
 CC fungicidal and cytostatic activities. (I) are useful, in human or
 CC veterinary medicine, in pharmaceutical compositions (for treating immune
 CC disorders, e.g. immune deficiency, autoimmune states, hypersensitivity,
 CC allergic graft rejection, infection, hormonal disorders and central
 CC nervous system diseases), also, where (I) is a MHC ligand (Ia), in
 CC vaccines for treatment or prevention of: (i) viral, bacterial, parasitic
 CC or fungal infections; or (ii) of cancers. A particular application is in
 CC anti-melanoma vaccines. (I) are also useful for in vitro diagnosis of
 CC diseases associated with interactions between MHC and (I), e.g. melanoma
 CC and human immunodeficiency virus infection. AAM98898 to AAM99592
 CC represent peptides which can be used in pharmaceutical compounds from
 CC the present invention.

XX Sequence 16 AA;

Query Match 30.4%; Score 31; DB 22; Length 16;

Best Local Similarity 75.0%; Pred. No. 2.6e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHOI 11

Db 8 EAEIAHQI 15

RESULT 15

AAW93362

ID AAW93362 standard; Protein; 18 AA.

AC AAW93362;

XX 28-MAY-1999 (first entry)

XX Human DPD protein fragment #1.

XX DPD; dihydropyrimidine-dehydrogenase; monoclonal antibody; MAB;
 KW immunoassay reagent; cancer patient; treatment; antitumor agent;
 KW 5-fluorouracil; affinity purification; toxicity.

OS Homo sapiens.

XX DE19837391-A1.

XX 25-FEB-1999.

XX 18-AUG-1998; 98DE-1037391.

XX 22-AUG-1997; 97EP-0114630.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Hasegawa M, Yoshikubo T;

XX WPI; 1999-155202/14.

XX Monoclonal antibody specific for dihydropyrimidine dehydrogenase -
 PT for assessing patient response to 5-fluorouracil antitumor agents

XX Disclosure; Page 24; 34pp; German.

XX This invention describes a monoclonal antibody (MAB) specific for
 CC dehydropyrimidine dehydrogenase (DPD). This MAB is used as immunoassay
 CC reagents to identify a lack of DPD in a patient and to assess the
 CC sensitivity of cancer patients to treatment with antitumor agents of the
 CC 5-fluorouracil (5-FU) type. The MAB can also be used for affinity
 CC purification of DPD. DPD is involved in reduction of 5-FU (and related
 CC catabolites and derivatives) and lack of it is associated with increased
 CC toxicity of this type of antitumor agent. It has specific binding
 CC interaction. The MAB provide a sensitive and reliable test for DPD,
 CC which is simple, rapid and suitable for routine screening.

XX Sequence 18 AA;

Query Match 30.4%; Score 31; DB 20; Length 18;

Best Local Similarity 50.0%; Pred. No. 2.9e+02;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 8 CHQIEYFGD 17

Db 1 CEKLENNFGD 10

Search completed: April 23, 2003, 13:42:57

Job time : 28.2022 secs

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:21:27 ; Search time 27.9101 Seconds
(without alignments)
132.885 Million cell updates/sec

Title: US-09-836-073-3
Perfect score: 98
Sequence: 1 AALQAKICHQIYFQGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL21.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mnc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	84.7	381	11 Q9CYB9	Q9cyb9 mus musculus
2	62	63.3	206	13 Q8QH15	Q8qh15 gallus gall
3	47	48.0	671	4 Q9UL65	Q9ul65 homo sapien
4	46	46.9	391	2 P72382	P72382 staphylococ
5	46	46.9	391	2 P95709	P95709 staphylococ
6	46	46.9	391	16 Q99X57	Q99x57 staphylococ
7	46	46.9	608	11 Q9QUG9	Q9qug9 mus musculus
8	45	45.9	568	16 Q9KL17	Q9kl17 vibrio chol
9	45	45.9	928	10 Q9LU02	Q9lu02 oryza sativ
10	44	44.9	239	10 Q9SMD8	Q9smd8 laminaria d
11	43	43.9	337	2 Q8RK98	Q8rk98 mycoplasma
12	42.5	43.4	577	16 Q97DB6	Q97db6 clostridium
13	42	42.9	396	5 O01806	O01806 caenorhabdi
14	42	42.9	541	16 Q92KY5	Q92ky5 helicobacte
15	42	42.9	542	16 Q25334	Q25334 helicobacte
16	42	42.9	658	10 Q9CAN3	Q9can3 arabidopsis

17	42	42.9	1220	12 Q39272	Q39272 equine herp
18	42	42.9	1249	12 Q65152	Q65152 african swi
19	42	42.9	2644	4 Q13535	Q13535 homo sapien
20	41	41.8	259	3 Q9P5X3	Q9p5x3 neurospora
21	41	41.8	296	5 Q9W129	Q9w129 drosophila
22	41	41.8	336	2 Q50359	Q50359 mycoplasma
23	41	41.8	336	16 Q98QD2	Q98qd2 mycoplasma
24	41	41.8	366	2 Q85797	Q85797 mycoplasma
25	41	41.8	369	2 Q30382	Q30382 mycoplasma
26	41	41.8	369	2 Q9R8C2	Q9r8c2 mycoplasma
27	41	41.8	369	16 Q98PP1	Q98pp1 mycoplasma
28	41	41.8	399	2 Q85799	Q85799 mycoplasma
29	41	41.8	411	10 Q9FL36	Q9fl36 arabidopsis
30	41	41.8	422	10 Q94A38	Q94a38 arabidopsis
31	41	41.8	434	5 Q961H8	Q961h8 drosophila
32	41	41.8	628	5 Q9V761	Q9v761 drosophila
33	40	40.8	106	11 Q9CPS8	Q9cps8 mus musculu
34	40	40.8	120	16 Q92ER1	Q92er1 listeria in
35	40	40.8	132	16 Q926P7	Q926p7 listeria in
36	40	40.8	133	16 Q9K9F1	Q9k9f1 bacillus ha
37	40	40.8	165	5 Q9GV84	Q9gv84 plasmodium
38	40	40.8	239	3 Q9P8N6	Q9p8n6 cochliobolu
39	40	40.8	453	10 Q8RWR2	Q8rwr2 arabidopsis
40	40	40.8	523	10 Q94K80	Q94k80 arabidopsis
41	40	40.8	532	2 Q85154	Q85154 photorhabdu
42	40	40.8	567	5 Q16108	Q16108 molgula occ
43	40	40.8	568	5 Q76785	Q76785 molgula occ
44	40	40.8	596	5 Q8T6A8	Q8t6a8 caenorhabdi
45	40	40.8	598	10 Q49535	Q49535 arabidopsis

ALIGNMENTS

RESULT 1

Q9CYB9 PRELIMINARY; PRT; 381 AA.
AC Q9CYB9;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Sjogren syndrome antigen B.
GN SSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Bojelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Togo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK017822; BAB30957.1; -
MGD; MGI:98423; SSB.

DR InterPro: IPR002344; Lupus_La.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; rim; 1.
 DR PRINTS: PR00302; LUPUSLA.
 SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;

Query Match 84.7%; Score 83; DB 11; Length 381;
 Best Local Similarity 82.4%; Pred. No. 4e-06; 1; Indels 0; Gaps 0;
 Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALQAKICHQIQYIFGQF 18
 ||:|||||:|||||
 Db 12 ALEAKICHQIEYIFGDF 28
 ||:|||||:|||||

RESULT 2
 Q8QHI5 PRELIMINARY; PRT; 206 AA.
 ID Q8QHI5;
 AC Q8QHI5;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Untranslated region binding-protein.
 GN UBP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA L'Ecuquer T.J., Fang H.-L.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF467897; AAL76269.1;
 SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFFB90E9 CRC64;

Query Match 63.3%; Score 62; DB 13; Length 206;
 Best Local Similarity 71.4%; Pred. No. 0.0089;
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LQAKICHQIQYIFG 16
 |:|:|:|:|:|:|
 Db 13 LESKICQIEYIFG 26
 |:|:|:|:|:|:|

RESULT 3
 Q9UL65 PRELIMINARY; PRT; 671 AA.
 ID Q9UL65;
 AC Q9UL65; 000538;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Guanine exchange factor MCG7 isoform 1 (F25B3.3 Kinase like protein)
 DE CALDAG-GEFI.
 GN Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20493616; PubMed-10918068;
 RA Clyde-Smith J., Silins G., Gartside M., Grimmond S., Etheridge M.,
 RA Apolloni A., Hayward N., Hancock J.F.;
 RT "Characterization of RasGRP2, a Plasma Membrane-targeted, Dual
 RT Specificity Ras/Rap Exchange Factor";
 RL J. Biol. Chem. 275:32260-32267(2000).
 RN [2]
 RP SEQUENCE OF 63-671 FROM N.A.
 RX MEDLINE-98001089; PubMed-9341881;
 RA Kedra D., Seroussi E., Fransson I., Trifunovic J., Clark M.,
 RA Lagercrantz J., Blennow E., Mehlin H., Dumaniski J.;
 RT "The germinal center kinase gene and a novel CDC25-like gene are

RT located in the vicinity of the PYGM gene on 11q13.";
 RL Hum. Genet. 100:611-619(1997).
 RN [3]
 RP SEQUENCE OF 63-671 FROM N.A.
 RX MEDLINE-99007305; PubMed-9789079;
 RA Kawasaki H., Springett G.M., Toki S., Canales J.J., Harlan P.,
 RA Blumenstiel J.P., Chen E.J., Bany I.A., Mochizuki N., Ashbacher A.,
 RA Matsuda M., Housman D.E., Graybiel A.M.;
 RT "A Rap guanine nucleotide exchange factor enriched highly in the basal
 RT ganglia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:13278-13283(1998).
 RN [4]
 RP SEQUENCE OF 63-671 FROM N.A.
 RA Kawasaki H., Housman D.E., Graybiel A.M.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 63-671 FROM N.A.
 RA Kawasaki H., Housman D.E., Graybiel A.M.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 63-671 FROM N.A.
 RA Kawasaki H., Springett G.M., Toki S., Canales J.J., Blumenstiel J.P.,
 RA Chen E.J., Bany I.A., Mochizuki N., Ashbacher A., Matsuda M.,
 RA Housman D.E., Graybiel A.M.;
 RT "A Novel Rap Guanine Nucleotide Exchange Factor Enriched in the Basal
 RT Ganglia.";
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 63-671 FROM N.A.
 RA Silins G.U., Grimmond S., Hayward N.;
 RT "Characterisation of a Novel Nucleotide Exchange Factor.";
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF043722; AAF07219.1;
 DR EMBL: Y12336; CAA73005.1;
 DR EMBL: U78170; AAD12741.1;
 DR EMBL: AF081194; AAC79698.1;
 DR EMBL: AF043723; AAF07220.1;
 DR HSP; P28867; IptQ.
 DR InterPro: IPR002219; DAG_PE-bind.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR000651; RasGEFN.
 DR InterPro: IPR001895; RasGRF_CDC25.
 DR Pfam: PF001130; DAG_PE-bind; 1.
 DR Pfam: PF00036; ehand; 2.
 DR Pfam: PF00617; RasGEF; 1.
 DR SMART; SM00109; C1; 1.
 DR SMART; SM00147; RasGEF; 1.
 DR SMART; SM00229; RasGEFN; 1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN_2.
 KW Phorbol-ester binding.
 SQ SEQUENCE 671 AA; 75547 MW; 67B7BD2B4F4AED4D CRC64;

Query Match 48.0%; Score 47; DB 4; Length 671;
 Best Local Similarity 41.2%; Pred. No. 10;
 Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 ALQAKICHQIQYIFGQF 18
 |:|:|:|:|:|:|
 Db 132 SLQVKTCHLVRYWISAF 148
 |:|:|:|:|:|:|

RESULT 4
 P72382 PRELIMINARY; PRT; 391 AA.
 ID P72382;
 AC P72382;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Cap8P.
 GN CAP8P.
 OS Staphylococcus aureus.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BECKER;
 RX MEDLINE=96178981; PubMed=8606192;
 RA Sau S., Lee C.Y.; Foster T.J., Lee J.C.;
 RT "Cloning of type 8 capsule genes and analysis of gene clusters for the
 RT production of different capsular polysaccharides in Staphylococcus
 RT aureus";
 RL J. Bacteriol. 178:2118-2126(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BECKER;
 RX MEDLINE=97197525;
 RA Sau S., Sun J., Lee C.Y.;
 RT "Molecular characterization and transcriptional analysis of type 8
 RT capsule genes in Staphylococcus aureus";
 RL J. Bacteriol. 179:1614-1621(1997).
 DR ENBL; U73374; AAB49445.1; -;
 DR HSP; P27828; IF6D.
 DR InterPro; IPR003331; Epimerase_2.
 DR Pfam; PF02350; Epimerase_2; 1.
 DR TIGRFAMS; TIGR00236; wecB; 1.
 SQ SEQUENCE 391 AA; 44247 MW; 1E8D9FAA9BC76F0D CRC64;
 Query Match 46.9%; Score 46; DB 2; Length 391;
 Best Local Similarity 63.6%; Pred. No. 9.2;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 6 KICHIQIYFG 16
 Db 366 RICEAIEYFG 376
 :||| :|||
 RESULT 5
 P95709 PRELIMINARY; PRT; 391 AA.
 ID P95709
 AC P95709;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Cap5P.
 GN CAP5P.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=REYNOLDS, AND NEWMAN;
 RX MEDLINE=97388587; PubMed=9245821;
 RA Sau S., Bhassin N., Wann E.R., Lee J.C., Foster T.J., Lee C.Y.;
 RT "The Staphylococcus aureus allelic genetic loci for serotype 5 and 8
 RT capsule expression contain the type-specific genes flanked by common
 RT genes";
 RL Microbiology 143:0-0(0).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=REYNOLDS, AND NEWMAN;
 RX MEDLINE=98101481; PubMed=9440531;
 RA Kiser K.B., Lee J.C.;
 RT "Staphylococcus aureus cap50 and cap5P genes functionally complement
 RT mutations affecting enterobacterial common-antigen biosynthesis in
 RT Escherichia coli.";
 RL J. Bacteriol. 180:403-406(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=REYNOLDS, AND NEWMAN;
 RX MEDLINE=98125727; PubMed=9466251;
 RA Bhassin N., Albus A., Michon F., Livolsi P.J., Park J.-S., Lee J.C.;
 RT "Identification of a gene essential for O-acetylation of the

RT Staphylococcus aureus type 5 capsular polysaccharide.";
 RL Mol. Microbiol. 27:9-21(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=REYNOLDS, AND NEWMAN;
 RX MEDLINE=96178981; PubMed=8606192;
 RA Bagga N., Wann E.R., Foster T.J., Lee J.C.;
 RT Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RL ENBL; U81973; AAC46099.1; -;
 DR HSP; P27828; IF6D.
 DR InterPro; IPR003331; Epimerase_2.
 DR Pfam; PF02350; Epimerase_2; 1.
 DR TIGRFAMS; TIGR00236; wecB; 1.
 SQ SEQUENCE 391 AA; 44328 MW; B0105F690DB7CF1D CRC64;
 Query Match 46.9%; Score 46; DB 2; Length 391;
 Best Local Similarity 63.6%; Pred. No. 9.2;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 6 KICHIQIYFG 16
 Db 366 RICEAIEYFG 376
 :||| :|||
 RESULT 6
 Q99X57 PRELIMINARY; PRT; 391 AA.
 ID Q99X57
 AC Q99X57;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Capsular polysaccharide synthesis enzyme Cap5P.
 GN CAPP OR SAV0164 OR SA0159.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
 OS Staphylococcus aureus (strain N315).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=158878, 158879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S. aureus (strain Mu50), and S. aureus (strain N315);
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kubara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 RT aureus";
 RL Lancet 357:1225-1240(2001).
 DR ENBL; AP003358; BAB56326.1; -;
 DR ENBL; AP003129; BAB41379.1; -;
 DR HSP; P27828; IF6D.
 DR InterPro; IPR003331; Epimerase_2.
 DR Pfam; PF02350; Epimerase_2; 1.
 DR TIGRFAMS; TIGR00236; wecB; 1.
 KW Complete proteome.
 SQ SEQUENCE 391 AA; 44372 MW; D0DF5FA715BCCECC CRC64;
 Query Match 46.9%; Score 46; DB 16; Length 391;
 Best Local Similarity 63.6%; Pred. No. 9.2;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 6 KICHIQIYFG 16
 Db 366 RICEAIEYFG 376
 :||| :|||
 RESULT 7
 Q90UG9 PRELIMINARY; PRT; 608 AA.
 ID Q90UG9
 AC Q90UG9;

RESULT 12
Q97DB6
ID Q97DB6
AC Q97DB6;
PRELIMINARY; PRT; 577 AA.

[2] SEQUENCE FROM N.A.
RP STRAIN-BRISTOL N2;
RC STRAIN-BRISTOL N2;
RA Sammons L., Wohldmann P., Gillam B.;
R. "The sequence of *C. elegans* cosmid C44E4.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
[3] SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

```

DR EMBL: AF003140; AAB54169.1; -.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm, 1.
DR PRINTS: PR00302; LOPUSLA.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS0102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
SQ SEQUENCE 396 AA; 43631 MW; A2D828A4FAA3C34 CRC64;

Query Match 42.9%; Score 42; DB 5; Length 396;
Best Local Similarity 63.6%; Pred. No. 45;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIOYFG 16
|||:||||
Db 15 KIIOLEYFG 25

RESULT 14
Q9ZKY5 PRELIMINARY; PRT; 541 AA.
AC Q9ZKY5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE PUTATIVE.
GN JHP0737.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
DR EMBL: AE001509; AAD06373.1; -.
KW Complete proteome.
SQ SEQUENCE 541 AA; 61715 MW; 3232E0A398B69B1E CRC64;

Query Match 42.9%; Score 42; DB 16; Length 541;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQ 12
|:|:||||:
Db 123 AEVQDKVCHQVE 134

RESULT 15
O25534 PRELIMINARY; PRT; 542 AA.
AC O25534;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein HP0863.
GN HP0863.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=2110;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;

```

```

RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
DR EMBL: AE000596; AAD07917.1; -.
DR TIGR: HP0863; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 542 AA; 61899 MW; E77A4EF68A04ECED CRC64;

Query Match 42.9%; Score 42; DB 16; Length 542;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQ 12
|:|:||||:
Db 123 AEVQDKVCHQVE 134

Search completed: April 23, 2003, 13:32:49
Job time : 29.9101 secs

```

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:42 ; Search time 4.95506 Seconds
(without alignments)
150.669 Million cell updates/sec

Title: US-09-836-073-3.

Perfect score: 98

Sequence: 1 AALQAKICHQIQYFGQF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	87	88.8	404	1 LA_BOVIN	P10881 bos taurus
2	87	88.8	408	1 LA_HUMAN	P03455 homo sapien
3	87	88.8	415	1 LA_RAT	P38656 rattus norv
4	83	84.7	415	1 LA_MOUSE	P32067 mus musculu
5	59	60.2	427	1 LAB_XENLA	P28049 xenopus lae
6	55	56.1	428	1 LAA_XENLA	P28048 xenopus lae
7	43	43.9	383	1 LA_AEDAL	Q26457 aedes albop
8	43	43.9	602	1 EXSA_BUCAL	P57530 buchmera ap
9	42	42.9	482	1 IFT5_HUMAN	Q13325 homo sapien
10	42	42.9	1220	1 DPOL_HSVB	P28858 equine herp
11	41	41.8	646	1 YE14_SCHPO	Q13869 schizosacch
12	40	40.8	479	1 PRL2_ARATH	Q39190 arabidopsis
13	39.5	40.3	503	1 TCPT_VIBCH	P29480 vibrio chol
14	39	39.8	104	1 CY2_RHOVA	P00082 rhodocicrob
15	39	39.8	183	1 ZEB2_WAIZE	P08031 zea mays (m
16	39	39.8	390	1 LA_DROME	Q05112 human papil
17	39	39.8	631	1 VE1_HPV30	Q05112 human papil
18	39	39.8	735	1 YD07_SCHPO	Q10432 schizosacch
19	39	39.8	926	1 ME19_DROME	Q24087 drosophila
20	39	39.8	1221	1 V143_NPVAC	P24307 autographa
21	38.5	39.3	940	1 ADA_DROME	P19126 drosophila
22	38.5	39.3	977	1 A2A1_HUMAN	Q95782 h adapter-r
23	38.5	39.3	977	1 A2A1_MOUSE	P17426 m adapter-r
24	38	38.8	251	1 FOL2_MOUSE	Q05685 mus musculu
25	38	38.8	305	1 HEM6_VIBCH	Q9kvt4 vibrio chol
26	38	38.8	786	1 ST5B_HUMAN	P51692 homo sapien
27	38	38.8	1088	1 PIGO_HUMAN	Q48teq homo sapien
28	38	38.8	1107	1 MYIA_MOUSE	P46735 mus musculu
29	38	38.8	1136	1 MYIA_RAT	Q05096 rattus norv
30	38	38.8	1181	1 HAIR_RAT	P97609 rattus norv
31	38	38.8	1189	1 HAIR_MOUSE	Q43593 homo sapien
32	38	38.8	1596	1 GLI3_MOUSE	Q61602 mus musculu
33	37	37.8	227	1 MOAR_KLEAE	P54794 klebsiella

34	37	37.8	231	1 MTN_BACSU	O32028 bacillus su
35	37	37.8	243	1 SDGF_RAT	P24338 rattus norv
36	37	37.8	359	1 KKR3_MOUSE	P51678 mus musculu
37	37	37.8	463	1 Y863_SYNY3	P73754 synechocyst
38	37	37.8	464	1 CASB_KLEOX	Q48409 klebsiella
39	37	37.8	466	1 SRO9_YEAST	P25567 saccharomyc
40	37	37.8	480	1 GLGA_RHITR	Q9euts rhizobium t
41	37	37.8	697	1 HRPI_ERWAM	P36554 erwinia amy
42	37	37.8	715	1 ORC3_MOUSE	Q91k30 mus musculu
43	37	37.8	719	1 FRE4_YEAST	P53746 saccharomyc
44	37	37.8	745	1 FSP1_RAT	Q63517 rattus norv
45	37	37.8	851	1 NCL1_CAEEL	P34611 caenorhabdi

ALIGNMENTS

RESULT 1					
LA_BOVIN					
ID	LA_BOVIN	STANDARD:	PRT:	404 AA.	
AC	P10881;				
DT	01-JUL-1989 (Rel. 11, Created)				
DT	01-JUL-1989 (Rel. 11, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).				
DE	SSB.				
GN	Bos taurus (Bovine).				
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OX	NCBI_TaxID=9913;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Pituitary;				
RX	MEDLINE=89202037; PubMed=2468131;				
RA	Chan E.K.L., Sullivan K.F., Tan E.M.;				
RT	"Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences for RNA-binding."				
RL	Nucleic Acids Res. 17:2233-2244(1989).				
CC	-1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S, AND 7-2 RNAs.				
CC	-1- SUBCELLULAR LOCATION: Nuclear (Probable).				
CC	-1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE C-TERMINAL PART OF THE PROTEIN.				
CC	-1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).				

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X13698; CAA31986.1; -
PIR; S03849; S03849.
InterPro; IPR002344; Lupus_La.
InterPro; IPR000504; RNA_rec_mot.
Pfam; PF00076; rrm; 1.
PRINTS; PR00302; LUPUSLA.
SMART; SM00360; RRM; 1.
PROSITE; PS00102; RRM; 1.
PROSITE; PS00030; RRM_RNP_1; 1.
RNA-binding; Nuclear protein; Phosphorylation.
DOMAIN 111 197
SEQUENCE 404 AA; 46534 MW; 4EE30B5C262AD6A1 CRC64;
Query Match 88.8%; Score 87; DB 1; Length 404;

Best Local Similarity 83.3%; Pred. No. 3.1e-07;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYVFGQF 18
DB 11 AALEAKICHQIEYVFGDF 28

RESULT 2
LA_HUMAN
ID LA_HUMAN STANDARD; PRT; 408 AA.
AC P03455;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lupus La protein (Sjogren syndrome type B antigen (SS-B)) (La
DE ribonucleoprotein) (La autoantigen).
GN SSB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89202037; PubMed=2468131;
RA Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
RT sequences for RNA-binding.";
RL Nucleic Acids Res. 17:2233-2244(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89053970; PubMed=3192525;
RA Chambers J.C., Kenan D., Martin B.J., Keene J.D.;
RT "Genomic structure and amino acid sequence domains of the human La
RT autoantigen.";
RL J. Biol. Chem. 263:18043-18051(1988).
RN [3]
RP SEQUENCE OF 54-408 FROM N.A.
RX MEDLINE=88199081; PubMed=2452201;
RA Sturgess A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
RA Coppel R.S.;
RT "Characteristics and epitope mapping of a cloned human autoantigen
RT La.";
RL J. Immunol. 140:3212-3218(1988).
RN [4]
RP SEQUENCE OF 54-97 FROM N.A.
RX MEDLINE=85166283; PubMed=3856888;
RA Chambers J.C., Keene J.D.;
RT "Isolation and analysis of cDNA clones expressing human lupus La
RT antigen.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119(1985).
RN [5]
RP FUNCTION.
RX MEDLINE=89251617; PubMed=2470590;
RA Gottlieb E., Steitz J.A.;
RT "Function of the mammalian La protein: evidence for its action in
RT transcription termination by RNA polymerase III.";
RL EMBO J. 8:851-861(1989).
RN [6]
RP PHOSPHORYLATION.
RX MEDLINE=97207017; PubMed=9054510;
RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Marate R.J.;
RT "Phosphorylation of the human La antigen on serine 366 can regulate
RT recycling of RNA polymerase III transcription complexes.";
RL Cell 88:707-715(1997).
RN [7]
CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III INCLUDING TRNA AND 4.5S, 5S, 7S,
CC AND 7-2 RNAs.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE

C-TERMINAL PART OF THE PROTEIN.
-1- PTM: THE N-TERMINUS IS BLOCKED.
-1- DISEASE: SERA FROM PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS
OFTEN CONTAIN ANTIBODIES THAT REACT WITH THE NORMAL CELLULAR
LA PROTEIN AS IF THESE ANTIGEN WAS FOREIGN.
-1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; X13697; CAA31985.1; -
EMBL; J04205; AAA51885.1; -
PIR; A31888; A31888.
PIR; A22956; A22956.
PIR; A31273; A31273.
PIR; S03848; S03848.
PIR; S11013; S11013.
Genew; HGNC:113116; SSB.
MIM; 109090; -
InterPro; IPR002344; Lupus.La.
InterPro; IPR000504; RNA_rec_mot.
Pfam; PF00076; rrm; 1.
PRINTS; PR00302; LUPUSLA.
SMART; SM00360; RRM; 1.
PROSITE; PS01012; RRM; 1.
PROSITE; PS00030; RRM_RNP_1; 1.
Systemic lupus erythematosus; RNA-binding; Phosphorylation;
KW Nuclear protein.
FT DOMAIN 111 187 RNA-BINDING (RRM).
FT MOD_RES 366 366 PHOSPHORYLATION (BY CK2).
SQ SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;
Query Match 88.8%; Score 87; DB 1; Length 408;
Best Local Similarity 83.3%; Pred. No. 3.2e-07;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYVFGQF 18
DB 11 AALEAKICHQIEYVFGDF 28

RESULT 3
LA_RAT
ID LA_RAT STANDARD; PRT; 415 AA.
AC P38656;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
DE homolog).
GN SSB OR SS-B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9324255; PubMed=7916708;
RA Senzel I., Troester H., Bartsch H., Schwemmler M., Igloi G.L.,
RA Bachmann M.;
RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La:
RT detection of species-specific variations.";
RL Gene 126:265-268(1993).
CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III INCLUDING TRNA AND 4.5S, 5S, 7S,
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,

CC AND 7-2 RNAs.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X67859; CAA48043.1; -
DR FIC: JCI1494; JCI1494.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
KW RNA-binding; Nuclear protein; Phosphorylation.
FT DOMAIN 111 187 RNA-BINDING (RRM).
SQ SEQUENCE 415 AA; 47777 MW; 033FD9CC1E475F98 CRC64;

Query Match 88.8%; Score 87; DB 1; Length 415;
Best Local Similarity 83.3%; Pred. No. 3.2e-07;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALQAKICHQIQYFGQF 18
DB 11 AALEAKICHQIQYFGDF 28

RESULT 4
LA_MOUSE
ID LA_MOUSE STANDARD; PRT; 415 AA.
AC P32067;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
DE homolog).
DE SSB OR SS-B.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93203630; PubMed=8454877;
RA Topfer F., Gordon T., McCluskey J.;
RT "Characterization of the mouse autoantigen La (SS-B). Identification
RT of conserved RNA-binding motifs, a putative ATP binding site and
RT reactivity of recombinant protein with poly(U) and human
RT autoantibodies.";
RL J. Immunol. 150:3091-3100(1993).
RN [2]
RP SEQUENCE OF 1-11 FROM N.A.
RA Groelz D., Bachmann M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
CC AND 7-2 RNAs.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L00993; AAA39415.1; -
DR EMBL: Y07951; CAA69249.1; -
DR MGD: MGI:98423; Ssb.
DR InterPro: IPR002344; Lupus_La.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; 1.
KW RNA-binding; Nuclear protein; Phosphorylation.
FT DOMAIN 111 187 RNA-BINDING (RRM).
SQ SEQUENCE 415 AA; 47756 MW; 2D75197692FDC933 CRC64;

Query Match 84.7%; Score 83; DB 1; Length 415;
Best Local Similarity 82.4%; Pred. No. 1.5e-06;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALQAKICHQIQYFGQF 18
DB 12 ALEAKICHQIQYFGDF 28

RESULT 5
LAB_XENLA
ID LAB_XENLA STANDARD; PRT; 427 AA.
AC P28049;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen
DE homolog B).
DE LAB1.
GN Xenopus laevis (African clawed frog).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93287095; PubMed=8510143;
RA Scherly D., Stutz F., Lin-Marq N., Clarkson S.G.;
RT "La proteins from xenopus laevis. cDNA cloning and developmental
RT expression.";
RL J. Mol. Biol. 231:196-204(1993).
CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- DEVELOPMENTAL STAGE: BARELY DETECTABLE IN STAGE I/II OOCYTES,
CC ACCUMULATE IN STAGE III/IV OOCYTES, THEN EXHIBIT A ROUGHLY
CC CONSTANT STAGE LEVEL IN MATURE OOCYTES, EGGS, AND EARLY
CC EMBRYOS.
CC -1- PTM: PHOSPHORYLATED (PROBABLE).
CC -1- MISCELLANEOUS: THERE ARE TWO FORMS OF LA, LAA AND LAB, IN XENOPUS.
CC -1- SIMILARITY: HIGH TO MAMMALIAN LA PROTEIN.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

DR PIR; S33818; S33818.
DR InterPro; IPR002344; Lupus_Ia.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF000076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS0030; RRM_RNP_1; 1.
DR RNA-binding; Nuclear protein; Phosphorylation.
FT DOMAIN 111 203 RNA-BINDING (RRM).
KW FT DOMAIN 196 212 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT FT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 428 AA; 48864 MW; AEB3A38B7D2E3EC3 CRC64;
SQ SEQUENCE 428 AA; 48864 MW; AEB3A38B7D2E3EC3 CRC64;

Query Match 56.1%; Score 55; DB 1; Length 428;
Best Local Similarity 81.8%; Pred. No. 0.077;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 KICHQIQYFG 16
   III II:IIII
DB 16 KICEQIEYFG 26

RESULT 7
LA_AEDAL STANDARD; PRT; 383 AA.
AC Q26457;
ID LA_AEDAL
DT 15-JUL-1998 (Rel. 36, Created)
DT 13-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
OS Aedes albopictus (Forest day mosquito).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Culicoidae; Aedes.
OC NCBI_TaxID=7160;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96113523; PubMed=8551578;
RA Pardigon N., Strauss J.H.;
RT "Mosquito homolog of the La autoantigen binds to Sindbis virus RNA.";
RL J. Virol. 70:1173-1181(1996).
CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTION TERMINATION BY RNA
CC POLYMERASE III. BINDS RNA AND DNA. BINDS TO THE 3' END OF THE
CC MINUS STRAND OF SINDBIS VIRUS RNA. THIS MAY BE SIGNIFICANT FOR
CC SINDBIS VIRUS RNA REPLICATION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. PRIMARILY NUCLEAR,, BUT SIGNIFICANT
CC AMOUNTS ARE PRESENT IN THE CYTOPLASM.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S80954; AAB35931.1;
CC InterPro; IPR002344; Lupus_La.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF000076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS0102; RRM; 1.
CC PROSITE; PS0030; RRM_RNP_1; FALSE_NEG.
CC RNA-binding; Nuclear protein; DNA-binding.
FT DOMAIN 141 228 RNA-BINDING (RRM).
FT SEQUENCE 383 AA; 44430 MW; 4E5CC8F21C40F452 CRC64;

Query Match 43.9%; Score 43; DB 1; Length 383;
Best Local Similarity 50.0%; Pred.No. 7.2;

```

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LOAKICHQIYFG 16
I:|:|:|:|
Db 43 LEASTIRQLEYFG 56

RESULT 8
EXSA_BUCAI STANDARD; PRT; 602 AA.
ID EX5A_BUCAI
AC P57530;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Exodeoxyribonuclease V alpha chain (EC 3.1.11.5).
GN RECD OR BU455
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RA MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS.";
RL Nature 407:81-86(2000).
CC -1- FUNCTION: EXHIBITS SEVERAL CATALYTIC ACTIVITIES, INCLUDING
CC ATP-DEPENDENT EXONUCLEASE, ATP-STIMULATED ENDONUCLEASE,
CC ATP-DEPENDENT UNWINDING AND DNA-DEPENDENT ATPASE ACTIVITIES.
CC STRAND CLEAVAGE OCCURS 5' TO 3' DURING THE UNWINDING OF DUPLEX
CC DNA AT CHI SEQUENCES, WHICH LOCALLY STIMULATE RECOMBINATION
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage (in the presence of
CC ATP) in either 5'- to 3'- or 3'- to 5'-direction to yield 5'-
CC phosphooligonucleotides.
CC -1- SUBUNIT: CONSIST OF THREE SUBUNITS; RECB, RECC AND RECD
CC (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001119; BAB13153.1; -
KW Hydrolyase; Nuclease; Exonuclease; Endonuclease; Helicase; ATP-binding;
KW DNA repair; Complete proteome.
FT NP_BIND 171 178 ATP (POTENTIAL).
SQ SEQUENCE 602 AA; 69494 MW; 340FFAE9BB436059 CRC64;

Query Match 43.9%; Score 43; DB 1; Length 602;
Best Local Similarity 47.1%; Pred. NO. 11;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 ALQAKICHQIYFGQF 18
I:|:|:|:|:|:|
Db 8 AVKLIRPIDFYSQF 24

RESULT 9
IFT5_HUMAN STANDARD; PRT; 482 AA.
ID IFT5_HUMAN
AC Q13325;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Retinolic acid- and interferon-inducible 58 kDa protein.
GN R156.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98114590; PubMed=9398535;
RA Nikiura T., Hirata R., Weil S.C.;
RT "A novel interferon-inducible gene expressed during myeloid
RT differentiation.";
RL Blood Cells Mol. Dis. 23:337-349(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas, and Spleen;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- INDUCTION: BY INTERFERONS.
CC -1- SIMILARITY: BELONGS TO THE IFIT FAMILY.
CC -1- SIMILARITY: CONTAINS 8 TPR REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U34605; AAA84934.1; -
DR EMBL; BC025786; AAH25786.1; -
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR; 5.
DR SMART; SM00028; TPR; 5.
KW Repeat; TPR repeat; Interferon induction.
FT REPEAT 51 84 TPR 1.
FT REPEAT 94 127 TPR 2.
FT REPEAT 138 173 TPR 3.
FT REPEAT 181 214 TPR 4.
FT REPEAT 249 282 TPR 5.
FT REPEAT 338 371 TPR 6.
FT REPEAT 376 410 TPR 7.
FT REPEAT 435 468 TPR 8.
SQ SEQUENCE 482 AA; 55846 MW; 80459C100384BE05 CRC64;

Query Match 42.9%; Score 42; DB 1; Length 482;
Best Local Similarity 60.0%; Pred. NO. 13;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 9 HQIQYFGQF 18
I:|:|:|:|:|
Db 376 HQIHYGRF 385

RESULT 10
DPOL_HSVB STANDARD; PRT; 1220 AA.
ID DPOL_HSVB
AC P28858;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7).
GN 30.
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92295566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1.";
RL Virology 189:304-316(1992).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + [DNA](N).

```

CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M86664; AAB02465.1; .
DR PIR; D36798; DJBEC3.
DR InterPro; IPR002064; DNA_pol_B.
DR Pfam; PF00136; DNA_pol_B; 1.
DR Pfam; PF03104; DNA_pol_B_exo; 1.
DR PRINTS; PR00106; DNAPOLB.
DR SMART; SM00486; POLBG; 1.
DR TRANSFERASE; P800116; DNA_POLYMERASE_B; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 1220 AA; 135956 MW; 858C14DCC071A5B CRC64;

Query Match 42.9%; Score 42; DB 1; Length 1220;
Best Local Similarity 60.0%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 ICHQIQYFG 16
   ||| ||||
DB 248 VCHTLYFG 257

RESULT 11
YE14_SCHPO
ID YE14_SCHPO STANDARD; PRT; 646 AA.
AC O13869;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative GTP-binding protein C1B3.04c.
GN SPAC1B3.04c.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer C., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

```

```

RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC -----
CC LEPA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z98598; CAB11233.1; .
DR HSP; F13551; LFNM.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR00795; EF_GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00009; Gtp_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR TIGRFAHS; TIGR00231; Small_GTP; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Hypothetical protein; GTP-binding.
FT NP_BIND 65 72 GTP (POTENTIAL).
FT NP_BIND 130 134 GTP (POTENTIAL).
FT NP_BIND 184 187 GTP (POTENTIAL).
FT SEQUENCE 646 AA; 72683 MW; F9FA9498D384503E CRC64;

Query Match 41.8%; Score 41; DB 1; Length 646;
Best Local Similarity 46.7%; Pred. No. 25;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 LOAKTCHQIQYFG 17
   ::|: | | | | |
DB 108 VKAQTCSMIYYHGQ 122

RESULT 12
PRL2_ARATH
ID PRL2_ARATH STANDARD; PRT; 479 AA.
AC Q39190; Q94007; Q9LUR9;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE PPI/PP2A phosphatases pleiotropic regulator PRL2.
GN PRL2 OR AT3G1650 OR MGL6.11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=98438452; PubMed=9765207;
RA Nemeth K., Saichert K., Putnoky P., Bhalerao R., Koncz-Kalman Z.,
RA Stankovic-Stangeland B., Bako L., Mathur J., Okresz L., Stabel S.,
RA Geigenberger F., Stitt M., Redei G.P., Schell J., Koncz C.;
RA "pleiotropic control of glucose and hormone responses by PRL1, a
RT nuclear WD protein, in Arabidopsis.";
RL Genes Dev. 12:3059-3073(1998).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT clones.";
RL DNA Res. 7:131-135(2000).
RN [3]
RP SEQUENCE FROM N.A.

```

```
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIEN Arabidopsis full length cDNA clones (RAPFs) sequenced by the
RL SSP consortium (Salk/Stanford/PGEC).";
CC -!- FUNCTION: PLIOTROPIC REGULATOR OF PP1 AND PP2A PHOSPHATASES.
CC -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: BELONGS TO THE PRL1/PRL2 FAMILY OF WD-REPEAT PROTEINS.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -!- CAUTION: It is uncertain whether Met-1 or Met-3 is the Initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X82826; CAA58033.1; -;
DR EMBL; AB022217; BAB02756.1; ALT_SEQ.
DR EMBL; AY054181; AAL06842.1; ALT_INIT.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD0000018; WD40; 4.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS00082; WD_REPEATS_2; 4.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT REPEAT 168 198 WD 1.
FT REPEAT 210 240 WD 2.
FT REPEAT 252 282 WD 3.
FT REPEAT 293 323 WD 4.
FT REPEAT 335 364 WD 5.
FT REPEAT 377 406 WD 6.
FT REPEAT 426 456 WD 7.
SQ SEQUENCE 479 AA; 53568 MW; 301986AAAEAR80670 CRC64;

Query Match 40.8%; Score 40; DB 1; Length 479;
Best Local Similarity 54.5%; Pred. No. 28;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIQYFG 16
Db 48 RLCHKIQVAFG 58
:|:|:| |

RESULT 13
TCPT_VIBCH STANDARD; PRT; 503 AA.
AC TCPT_VIBCH
AC P29480; Q9KTR0;
DT 01-APR-1993 (Rel. 25, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Toxin coregulated pilus biosynthesis protein T (TCP pilus
DE biosynthesis protein tcpt).
GN TCPT OR VC0835.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Classical Inaba 217561 / Serotype O1;
RX MEDLINE=93231537; PubMed=8097178;
RA Ogierman M.A., Zabihi S., Mourtzios L., Manning P.A.;
RT "Genetic organization and sequence of the promoter-distal region of
RT the top gene cluster of Vibrio cholerae.";
RL Gene 126:51-60(1993).
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN=cv. Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -!- FUNCTION: INVOLVED IN THE TRANSLLOCATION OF THE TCPA PILIN.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: BELONGS TO THE PULE/OUTE/EXEE/XPSE/XCPR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X64098; CAA45462.1; -;
DR EMBL; AE004168; AAF93998.1; -;
DR PIR; JN0524; JN0524.
DR TIGR; VC0835; -;
DR InterPro; IPR001482; GSPIL_E.
DR Pfam; PF00437; GSPIL_E; 1.
DR ProDom; PD000739; GSPIL_E; 1.
DR PROSITE; PS00662; T2SP_E; 1.
KW Transport; ATP-binding; Complete proteome.
FT NP_BIND 236 243 ATP (POTENTIAL).
FT CONFLICT 137 137 A -> S (IN REF. 1).
FT SEQUENCE 503 AA; 57276 MW; ED3FD8FED579F18 CRC64;

Query Match 40.3%; Score 39.5; DB 1; Length 503;
Best Local Similarity 43.8%; Pred. No. 36;
Matches 7; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 3 LOAKICHQ-IQYFQG 17
Db 471 LKCKVCHEFMHFGQ 486
:|:|:| |

RESULT 14
CY2_RHOVA STANDARD; PRT; 104 AA.
AC CY2_RHOVA
AC P00082;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Cytochrome C2.
OS Rhodocrobium vannielii.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Hyphomicrobium group; Rhodocrobium.
OX NCBI_TaxID=1069;
RN [1]
RP SEQUENCE.
RC STRAIN-ATCC 17100;
RX MEDLINE=76102814; PubMed=174109;
RA Ambler R.P., Meyer T.E., Kamen M.D.;
RT "Primary structure determination of two cytochromes c2: close
RT similarity to functionally unrelated mitochondrial cytochrome C.";
RL Proc. Natl. Acad. Sci. U.S.A. 73:472-475(1976).
CC -!- FUNCTION: CYTOCHROME C2 IS FOUND MAINLY IN PURPLE, NONSULFUR.
CC PHOTOSYNTHETIC BACTERIA WHERE IT FUNCTIONS AS THE ELECTRON DONOR
CC TO THE OXIDIZED BACTERIOCHLOROPHYLL IN THE PHOTOPHOSPHORYLATION
CC PATHWAY. HOWEVER, IT MAY ALSO HAVE A ROLE IN THE RESPIRATORY CHAIN
CC AND IS FOUND IN SOME NONPHOTOSYNTHETIC BACTERIA.
CC -!- PTM: BINDS ONE HEME GROUP PER MOLECULE.
DR PIR; A00074; CCRD2.
```

```

DR HSSP; P00083; ICYR.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR003088; Cyt_C1.
DR InterPro; IPR002327; Cyt_C1AB.
DR Pfam; PF00034; cytochrome_c; 1.
DR PRINTS; PR00604; CYTCHRMCIAB.
DR PRODOM; PD000375; Cyt_C1AB; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
KW Electron transport; Photosynthesis; Heme.
FT BINDING 14 14 HEME (COVALENT).
FT BINDING 17 17 HEME (COVALENT).
FT METAL 18 18 IRON (HEME AXIAL LIGAND).
FT METAL 80 80 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE 104 AA; 11209 MW; 1B5F2A7B657CD36F CRC64;

Query Match 39.8%; Score 39; DB 1; Length 104;
Best Local Similarity 75.0%; Pred. No. 9.9;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 QAKICHOI 11
| | | | |
DB 13 QKICHOV 20

DR Pfam; PF00234; tryp_alpha_amyl; 1.
DR PRINTS; PR00208; GLIADGLUTEN.
DR PRINTS; PR00211; GLUTELIN.
DR SMART; SM00499; AAI; 1.
KW Seed storage protein; Repeat; Multigene family; Signal.
FT SIGNAL 1 19
FT CHAIN 20 183 ZEIN-BETA.
SQ SEQUENCE 183 AA; 19558 MW; 3965BEBCL1151F45A CRC64;

Query Match 39.8%; Score 39; DB 1; Length 183;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALQAKICHQIQ 12
| | | : | | |
DB 108 ALQOCCCHQIR 118

Search completed: April 23, 2003, 13:28:05
Job time : 6.95506 secs

```

```

RESULT 15
ZEB2_MAIZE
ID ZEB2_MAIZE STANDARD; PRT; 183 AA.
AC P08031;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Zein-beta precursor (Zein 2) (16 kDa) (Zein 2C1).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87248094; PubMed=3596247;
RA Prat S., Perez-Grau L., Puigdomenech P.;
RT "Multiple variability in the sequence of a family of maize endosperm
RT proteins.";
RL Gene 52:41-49(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. W64A; TISSUE=Endosperm;
RX MEDLINE=91057131; PubMed=2243787;
RA Reina M., Guillen P., Ponte I., Boronat A., Palau J.;
RT "DNA sequence of the gene encoding the zcl protein from Zea mays W64
RT A.";
RL Nucleic Acids Res. 18:6425-6425(1990).
CC -1- FUNCTION: ZEINS ARE MAJOR SEED STORAGE PROTEINS.
CC -1- SUBCELLULAR LOCATION: ENDOSPERM PROTEIN BODIES.
CC -1- SIMILARITY: TO GLUTENIN 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16460; AAA33523.1; -
CC EMBL; X53515; CAA37595.1; -
CC PIR; B29017; B29017.
CC PIR; S12140; S12140.
CC MalzeDB; 58053; -
CC InterPro; IPR003612; AAI.
CC InterPro; IPR001954; Glia_glutenin.
CC InterPro; IPR000480; Glutelin.
CC InterPro; IPR001768; Try/amy1_inhtr.

```

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:03 ; Search time 10.1124 Seconds
(without alignments)
171.119 Million cell updates/sec

Title: US-09-836-073-3
Perfect score: 98
Sequence: 1 AALQAKICHQIQYFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-73.*

1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	88.8	404	1 S03849	ribonucleoprotein
2	87	88.8	408	1 A31888	ribonucleoprotein
3	87	88.8	415	1 JC1494	ribonucleoprotein
4	59	60.2	427	1 S33817	ribonucleoprotein
5	55	56.1	428	1 S33818	ribonucleoprotein
6	46	46.9	391	2 H89777	capsular polysacch
7	45	45.9	568	2 C82379	response regulator
8	43	43.9	602	2 A84983	exodeoxyribonuclea
9	42.5	43.4	577	2 D97337	mismatch repair pr
10	42	42.9	396	2 T30953	hypothetical prote
11	42	42.9	482	2 G02058	retinoic acid- and
12	42	42.9	541	2 H71887	hypothetical prote
13	42	42.9	542	2 G64627	hypothetical prote
14	42	42.9	658	2 D96656	hypothetical prote
15	42	42.9	1220	1 D7BEC3	DNA-directed DNA p
16	42	42.9	1220	2 T42573	DNA-directed DNA p
17	41	41.8	259	2 T49596	related to H-tran
18	41	41.8	336	2 S49396	HsdSB protein - M
19	41	41.8	336	2 B90566	restriction-modifi
20	41	41.8	369	2 G90596	restriction-modifi
21	41	41.8	645	2 T38022	probable GTP-bindi
22	40	40.8	120	2 AF1482	hypothetical prote
23	40	40.8	133	2 A83987	ABC transporter (p
24	40	40.8	175	2 S23687	erythrocyte membra
25	40	40.8	380	2 JC5217	site-specific DNA-
26	40	40.8	598	2 T05888	hypothetical prote
27	40	40.8	658	2 T19487	hypothetical prote
28	39.5	40.3	163	2 H82733	pathogenicity-rela
29	39.5	40.3	503	2 F82275	toxin co-regulated

30 39.5 40.3 503 2 JN0524
31 39 39.8 104 1 CCRD2
32 39 39.8 166 2 T32701
33 39 39.8 183 2 B29017
34 39 39.8 346 2 A82659
35 39 39.8 388 2 AF0183
36 39 39.8 390 2 A53773
37 39 39.8 390 2 A53781
38 39 39.8 461 2 T27651
39 39 39.8 487 2 T27653
40 39 39.8 505 2 B90181
41 39 39.8 506 2 F85016
42 39 39.8 529 2 T00677
43 39 39.8 631 2 S36505
44 39 39.8 696 2 F83886
45 39 39.8 735 2 T41187

ALIGNMENTS

RESULT 1

S03849

ribonucleoprotein La - bovine

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B

C:Species: Bos primigenius taurus (cattle)

C>Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: S03849

R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.

Nucleic Acids Res. 17, 2233-2244, 1989

A:Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequen

A:Reference number: S03848; MUID:89202037; PMID:2468131

A:Accession: S03849

A:Molecule type: mRNA

A:Residues: 1-404 <CHA>

A:Cross-references: EMBL:X13698; NID:9755; PIDN:CAA31986.1; PID:9756

A:Note: Part of this sequence was confirmed by protein sequencing

C:Comment: This protein associates with a variety of small RNA molecules, most of w
ay act as a transcription termination factor.

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: blocked amino end; phosphoprotein; RNA binding

F:113-178/Domain: ribonucleoprotein repeat homology <RMA>

F:113-118/Region: RNA-binding RNP2 motif

F:151-158/Region: RNA-binding RNP1 motif

F:228-404/Domain: phosphorylated #status predicted <PHY>

Query Match 88.8%; Score 87; DB 1; Length 404;

Best Local Similarity 83.3%; Pred. No. 3.6e-07;

Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 AALQAKICHQIQYFGQF 18

Db 11 AALEAKICHQIEYFGDF 28

RESULT 2

A31888

ribonucleoprotein La - human

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome an

C:Species: Homo sapiens (man)

C>Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: A31888; S03848; A22956; A61051; S11013; I55553; I70205; I70206; A31273

R:Chambers, J.C.; Kenan, D.; Martin, B.J.; Keene, J.D.

J. Biol. Chem. 263, 18043-18051, 1988

A:Title: Genomic structure and amino acid sequence domains of the human La autoanti

A:Reference number: A31888; MUID:89053970; PMID:3192525

A:Accession: A31888

A:Molecule type: mRNA

A:Residues: 1-408 <CHA>

A:Cross-references: GB:J04205; NID:9178686; PIDN:AAA51885.1; PID:9178687

R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.

Nucleic Acids Res. 17, 2233-2244, 1989

A:Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequen

JC1494

ribonucleoprotein La - rat

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B

C:Species: Rattus norvegicus (Norway rat)

C:Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: JC1494; S25145

R:Semsei, I.; Troester, H.; Bartsch, H.; Schwemmle, M.; Igloi, G.L.; Bachmann, M.

Gene 126, 265-268, 1993

A:Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection c

A:Reference number: JC1494; MUID:93246255; PMID:7916708

A:Accession: JC1494

A:Molecule type: mRNA

A:Residues: 1-415 <SEM>

A:Cross-references: GB:X67859; NID:g55778; PIDN:CAA48043.1; PID:g55779

A:Experimental source: liver

C:Comment: This protein associates with a variety of small RNA molecules, most of whi

ay act as a transcription termination factor.

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: phosphoprotein; RNA binding

F:112-178/Domain: ribonucleoprotein repeat homology <RRM>

F:113-118/Region: RNA-binding RNP2 motif

F:151-158/Region: RNA-binding RNP1 motif

F:227-415/Domain: phosphorylated #status predicted <PHY>

Query Match 88.8%; Score 87; DB 1; Length 415;

Best Local Similarity 83.3%; Pred. No. 3.7e-07;

Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHOIQYVGF 18

DB 11 AALEAKICHOIEYFGDF 28

|||||

RESULT 4

S33817

ribonucleoprotein La.B - African clawed frog

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B

C:Species: Xenopus laevis (African clawed frog)

C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: S33817; S28544

R:Scheryl, D.; Stutz, F.; Lin-Matq, N.; Clarkson, S.G.

J. Mol. Biol. 231, 196-204, 1993

A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression.

A:Reference number: S33817; MUID:93287095; PMID:8510143

A:Accession: S33817

A:Molecule type: mRNA

A:Residues: 1-427 <SCH>

A:Cross-references: EMBL:X68818; NID:g64875; PIDN:CAA48716.1; PID:g64876

C:Comment: This protein associates with a variety of small RNA molecules, most of whi

ay act as a transcription termination factor.

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: phosphoprotein; RNA binding

F:111-177/Domain: ribonucleoprotein repeat homology <RRM>

F:112-117/Region: RNA-binding RNP2 motif.

F:150-157/Region: RNA-binding RNP1 motif.

F:227-427/Domain: phosphorylated #status predicted <PHY>

Query Match 60.2%; Score 59; DB 1; Length 427;

Best Local Similarity 71.4%; Pred. No. 0.026;

Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LQAKICHOIQYVFG 16

DB 12 LDTKICGIEYVFG 25

|||||

RESULT 5

S33818

ribonucleoprotein La.A - African clawed frog

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B

C:Species: xenopus laevis (African clawed frog)

C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: S33818; S28545

R;Scherly, D.; Stutz, F.; Lin-Warq, N.; Clarkson, S.G.

J. Mol. Biol. 231, 196-204, 1993

A;Title: La proteins from *Xenopus laevis*. cDNA cloning and developmental expression.

A;Reference number: S33817; MUID:93287095; PMID:8510143

A;Accession: S33818

A;Molecule type: mRNA

A;Residues: 1-428 <SCH>

A;Cross-references: EMBL:X58817; NID:964873; PIDN:CAA48715.1; PID:964874

C;Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.

C;Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C;Keywords: phosphoprotein; RNA binding

F;112-178/Domain: ribonucleoprotein repeat homology <RRM>

F;113-118/Region: RNA-binding RNP2 motif

F;151-158/Region: RNA-binding RNP1 motif

F;228-428/Domain: phosphorylated #status predicted <PHY>

Query Match 56.1%; Score 55; DB 1; Length 428;

Best Local Similarity 81.8%; Pred. No. 0.13;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 6 KICHIQIYFG 16

Db 16 KICEQIEYFG 26

||| ||| |||

RESULT 6

H89777

capsular polysaccharide synthesis enzyme Cap5P [imported] - *Staphylococcus aureus* (strain

C;Species: *Staphylococcus aureus*

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C;Accession: H89777

R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogino,

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: H89777

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-391 <KUR>

A;Cross-references: GB:BA000018; PID:gl3700080; PIDN:BA41379.1; GSPDB:GN00149

A;Experimental source: strain N315

C;Genetics:

A;Gene: cap

C;Superfamily: lipopolysaccharide biosynthesis protein bplD

Query Match 46.9%; Score 46; DB 2; Length 391;

Best Local Similarity 63.6%; Pred. No. 4.3;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 6 KICHIQIYFG 16

Db 366 RICEAIEYFG 376

||| ||| |||

RESULT 7

C82379

response regulator VCAM086 [imported] - *Vibrio cholerae* (strain M16961 serogroup O1)

C;Species: *Vibrio cholerae*

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C;Accession: C82379

R;Heidelberg, J.F.; Elsen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: C82379

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-568 <HEI>

A;Cross-references: GB:AE004434; GB:AE003853; NID:9658531; PIDN:AAF96979.1; GSPDB: A;Experimental source: serogroup O1; strain M16961; biotype El Tor

C;Genetics:

A;Gene: VCAM086

A;Map position: 2

Query Match 45.9%; Score 45; DB 2; Length 568;

Best Local Similarity 42.9%; Pred. No. 9.1;

Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LOAKICHIQIYFG 16

Db 156 LMEEMCHQVEHIFG 169

||| ||| |||

RESULT 8

A84983

exodeoxyribonuclease V (EC 3.1.11.5) 67 kD polypeptide [imported] - *Buchnera* sp. (s

C;Species: *Buchnera* sp.

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001

C;Accession: A84983

R;Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A;Title: Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera*;

A;Reference number: A84930; MUID:20445173; PMID:10993077

A;Accession: A84983

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-602 <STO>

A;Cross-references: GB:AP000398; GSPDB:GN00144

A;Experimental source: strain APS

C;Genetics:

A;Gene: recB; BU455

C;Superfamily: exodeoxyribonuclease V 67K chain

C;Keywords: hydrolase

Query Match 43.9%; Score 43; DB 2; Length 602;

Best Local Similarity 47.1%; Pred. No. 21;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ALQAKICHIQIYFGQF 18

Db 8 AVKLKIRPIDYFYSQF 24

||| ||| |||

RESULT 9

D97337

mismatch repair protein Muts-like ATPase [imported] - *Clostridium acetobutylicum*

C;Species: *Clostridium acetobutylicum*

C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C;Accession: D97337

R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.;

J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing *Bacteriu*

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: D97337

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-577 <KUR>

A;Cross-references: GB:AE001437; PIDN:AAK81487.1; PID:gl5026659; GSPDB:GN00168

A;Experimental source: *Clostridium acetobutylicum* ATCC824

C;Genetics:

A;Gene: CAC3563

Query Match 43.4%; Score 42.5; DB 2; Length 577;

Best Local Similarity 38.1%; Pred. No. 25;

Matches 8; Conservative 6; Mismatches 2; Indels 5; Gaps 1;

Qy 3 LOAKICHIQIY-----YFGQF 18

Db 53 MKAKIHKLNLYERFKRYNGEF 73

||| ||| |||

RESULT 10

T30953

hypothetical protein C44E4.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000

C:Accession: T30953

R:Sammons, L.; Wohlmann, P.; Gillam, B.

submitted to the EMBL Data Library, August 1999

A:Description: The sequence of C. elegans cosmid C44E4.

A:Reference number: Z20945

A:Accession: T30953

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-396 <SAM>

A:Cross-references: EMBL:AF003140; PIDN:AA854169.1

A:Experimental source: strain Bristol N2; clone C44E4

C:Genetics:

A:Map position: I

A:Introns: 45/1; 114/3

A:Note: C44E4.4

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

Query Match 42.9%; Score 42; DB 2; Length 396;

Best Local Similarity 63.6%; Pred. No. 21;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIYFYG 16

||| ||:||||

Db 15 KIHKLEYFYG 25

RESULT 11

G02058

retinoic acid- and interferon-inducible 58K protein RI58 - human

C:Species: Homo sapiens (man)

C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 29-Sep-1999

C:Accession: G02058

R:Weil, S.C.

submitted to the EMBL Data Library, August 1995

A:Reference number: H00758

A:Accession: G02058

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-482 <WEI>

A:Cross-references: EMBL:U34605; NID:g1144510; PIDN:AA84934.1; PID:g1144511

C:Superfamily: Interferon-induced 56K protein

Query Match 42.9%; Score 42; DB 2; Length 482;

Best Local Similarity 60.0%; Pred. No. 26;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 9 HQIQYFQGF 18

||| ||:||||

Db 376 HQIHVYGRF 385

RESULT 12

H71887

hypothetical protein jhp0797 - Helicobacter pylori

C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 28-Jul-2000

C:Accession: H71887

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, E.D.; Doig, P.C.; Smith, D.R.;

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A:Reference number: A71800; MUID:99120557; PMID:9923662

A:Accession: H71887

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-541 <ARN>

A:Cross-references: GB:AE001509; GB:AE001439; NID:g4155350; PIDN:AA06373.1; PID:g41:

A:Experimental source: strain J99

C:Genetics:

A:Gene: jhp0797

C:Superfamily: Helicobacter pylori hypothetical protein jhp0797

Query Match 42.9%; Score 42; DB 2; Length 541;

Best Local Similarity 50.0%; Pred. No. 29;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHOIQ 12

| | | | |

Db 123 AEVQDKVCHQVE 134

RESULT 13

G64627

hypothetical protein HP0863 - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 28-Jul-2000

C:Accession: G64627

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann,

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McK

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9452185

A:Accession: G64627

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-542 <TOM>

A:Cross-references: GB:AE000596; GB:AE000511; NID:g2313982; PIDN:AA07917.1; PID:g23

C:Superfamily: Helicobacter pylori hypothetical protein jhp0797

Query Match 42.9%; Score 42; DB 2; Length 542;

Best Local Similarity 50.0%; Pred. No. 29;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHOIQ 12

| | | | |

Db 123 AEVQDKVCHQVE 134

RESULT 14

D96656

hypothetical protein F16M19.21 [Imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: D96656

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzi

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.B.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D96656

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-658 <STO>

A:Cross-references: GB:AE005173; NID:g6598840; PIDN:AAF18695.1; GSPDB:GN00141

C:Genetics:

A:Gene: F16M19.21

A:Map position: 1

Query Match 42.9%; Score 42; DB 2; Length 658;

Best Local Similarity 31.2%; Pred. No. 35;

Matches 5; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 3 LOAKICHOIQYFGQF 18
I: : : : :
Db 520 LETRCNSLKYYSAMF 535

RESULT 15

DJBEC3
DNA-directed DNA polymerase (EC 2.7.7.7) - equine herpesvirus 1 (strain Ab4p)
C;Species: equine herpesvirus 1
A;Note: host Equus caballus (domestic horse)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 11-Jun-1999
C;Accession: D36798
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
submitted to GenBank, March 1992
A;Description: The DNA sequence of equine herpesvirus-1.
A;Reference number: A36805
A;Accession: D36798
A;Molecule type: DNA
A;Residues: 1-1220 <TEL>
A;Cross-references: GB:M86664; NID:g330791; PIDN:AAB02465.1; PID:g330822
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A;Title: The DNA sequence of equine herpesvirus-1.
A;Reference number: A41831; MUID:92295566; PMID:1318606
A;Contents: annotation; possible protein-coding frames
A;Note: neither amino acid nor nucleotide sequence is given
C;Genetics:
A;Gene: 30
C;Superfamily: herpesvirus DNA-directed DNA polymerase
C;Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 42.9%; Score 42; DB 1; Length 1220;
Best Local Similarity 60.0%; Pred. NO. 63;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 ICHQIQYFG 16
: : : : :
Db 248 VCHTTLTYFG 257

Search completed: April 23, 2003, 13:34:34
Job time : 11.1124 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:27:28 ; Search time 10.3146 Seconds
(without alignments)
139.834 Million cell updates/sec

Title: US-09-836-073-3
Perfect score: 98
Sequence: 1 AALQAKICHQIYYFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	18	US-09-836-073-3	Sequence 3, Appli
2	87	88.8	18	US-09-836-073-1	Sequence 1, Appli
3	87	88.8	18	US-09-836-073-14	Sequence 14, Appli
4	87	88.8	460	US-10-102-806-695	Sequence 695, App
5	83	84.7	17	US-09-836-073-13	Sequence 13, Appli
6	81	82.7	18	US-09-836-073-9	Sequence 9, Appli
7	79	80.6	18	US-09-836-073-11	Sequence 11, Appli
8	79	80.6	18	US-09-836-073-12	Sequence 12, Appli
9	78	79.6	19	US-09-836-073-10	Sequence 10, Appli
10	76.5	78.1	19	US-09-836-073-16	Sequence 16, Appli
11	75	76.5	18	US-09-836-073-2	Sequence 2, Appli
12	75	76.5	18	US-09-836-073-4	Sequence 4, Appli
13	72	73.5	18	US-09-836-073-7	Sequence 7, Appli
14	71	72.4	18	US-09-836-073-8	Sequence 8, Appli
15	65	66.3	18	US-09-836-073-15	Sequence 15, Appli
16	63	64.3	18	US-09-836-073-5	Sequence 5, Appli
17	56	57.1	18	US-09-836-073-6	Sequence 6, Appli
18	50	51.0	38	US-09-843-676-25	Sequence 25, Appli
19	50	51.0	38	US-09-766-253-25	Sequence 25, Appli

20	50	51.0	38	9	US-09-438-486-25	Sequence 25, Appli
21	50	51.0	38	9	US-10-053-758-25	Sequence 25, Appli
22	50	51.0	38	9	US-10-054-295-25	Sequence 25, Appli
23	50	51.0	38	9	US-10-054-611-25	Sequence 25, Appli
24	49.5	50.5	37	9	US-09-843-676-24	Sequence 24, Appli
25	49.5	50.5	37	9	US-09-766-253-24	Sequence 24, Appli
26	49.5	50.5	37	9	US-09-438-486-24	Sequence 24, Appli
27	49.5	50.5	37	9	US-10-053-758-24	Sequence 24, Appli
28	49.5	50.5	37	9	US-10-054-295-24	Sequence 24, Appli
29	49.5	50.5	37	9	US-10-054-611-24	Sequence 24, Appli
30	45	45.9	16	9	US-09-836-073-19	Sequence 19, Appli
31	42	42.9	186	9	US-10-013-315-6	Sequence 6, Appli
32	41	41.8	569	10	US-09-925-300-1583	Sequence 1583, Ap
33	40	40.8	111	10	US-09-729-835-61	Sequence 61, Appli
34	40	40.8	2169	9	US-09-738-626-5455	Sequence 5455, Ap
35	39	39.8	18	9	US-09-836-073-17	Sequence 17, Appli
36	39	39.8	39	9	US-09-843-676-26	Sequence 26, Appli
37	39	39.8	39	9	US-09-766-253-26	Sequence 26, Appli
38	39	39.8	39	9	US-09-438-486-26	Sequence 26, Appli
39	39	39.8	39	9	US-10-053-758-26	Sequence 26, Appli
40	39	39.8	39	9	US-10-054-295-26	Sequence 26, Appli
41	39	39.8	39	9	US-10-054-611-26	Sequence 26, Appli
42	38.5	39.3	937	9	US-09-964-899-51	Sequence 51, Appli
43	38	38.8	984	9	US-10-024-368-2	Sequence 2, Appli
44	38	38.8	1089	9	US-10-174-590-266	Sequence 266, App
45	38	38.8	1089	9	US-10-176-758-266	Sequence 266, App

ALIGNMENTS

RESULT 1
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balgva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match 100.0%; Score 98; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AALQAKICHQIYYFGQF 18
Db 1 AALQAKICHQIYYFGQF 18

RESULT 2
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balgva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-1

Query Match 88.8%; Score 87; DB 9; Length 18;
 Best Local Similarity 83.3%; Pred. No. 6.7e-08;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYVFGQF 18
 |||:|||||:|||||
 Db 1 AALEAKICHQIEYVFGDF 18

RESULT 3

US-09-836-073-14
 ; Sequence 14, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Balidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 22002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Bovine
 US-09-836-073-14

Query Match 88.8%; Score 87; DB 9; Length 18;
 Best Local Similarity 83.3%; Pred. No. 6.7e-08;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYVFGQF 18
 |||:|||||:|||||
 Db 1 AALEAKICHQIEYVFGDF 18

RESULT 4

US-10-102-806-695
 ; Sequence 695, Application US/10102806
 ; Publication No. US20030054421A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA103PIC1
 ; CURRENT APPLICATION NUMBER: US/10/102,806
 ; CURRENT FILING DATE: 2002-03-22
 ; PRIOR APPLICATION NUMBER: 09/925,298
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05881
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 845
 ; SOFTWARE: PatentIn ver. 2.0
 ; SEQ ID NO 695
 ; LENGTH: 460
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-10-102-806-695

Query Match 88.8%; Score 87; DB 9; Length 460;
 Best Local Similarity 83.3%; Pred. No. 1.6e-06;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYVFGQF 18
 |||:|||||:|||||
 Db 63 AALEAKICHQIEYVFGDF 80

RESULT 5

US-09-836-073-13
 ; Sequence 13, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Balidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 22002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 17
 ; TYPE: PRT
 ; ORGANISM: Mouse
 US-09-836-073-13

Query Match 84.7%; Score 83; DB 9; Length 17;
 Best Local Similarity 82.4%; Pred. No. 2.9e-07;
 Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 AALQAKICHQIQYVFGQF 18
 |||:|||||:|||||
 Db 1 AALEAKICHQIEYVFGDF 17

RESULT 6

US-09-836-073-9
 ; Sequence 9, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Balidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 22002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-9

Query Match 82.7%; Score 81; DB 9; Length 18;
 Best Local Similarity 87.5%; Pred. No. 6.4e-07;
 Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYVFG 16
 |||:|||||:|||||
 Db 1 AALEAKICHQIEYVFG 16

```

RESULT 7
US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match      80.6%; Score 79; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 1.4e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AALQAKICHQIQYFGQF 18
   |||:|||||:|||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 8
US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match      80.6%; Score 79; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 1.4e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AALQAKICHQIQYFGQF 18
   |||:|||||:|||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 9
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match      79.6%; Score 78; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 2e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AALQAKICHQIQYFGQF 18
   |||:|||||:|||||
Db 1 AALEAKICHQIEYFGDF 18

```

```

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match      79.6%; Score 78; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 2e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AALQAKICHQIQYFGQF 18
   |||:|||||:|||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 10
US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match      78.1%; Score 76.5; DB 9; Length 19;
Best Local Similarity 78.9%; Pred. No. 3.7e-06;
Matches 15; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 AALQAKICHQIQYFGQF 19
   |||:|||||:|||||
Db 1 AALEAKICHQIEYFGDF 19

RESULT 11
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

```

Query Match 76.5%; Score 75; DB 9; Length 18;
Best Local Similarity 72.2%; Pred. No. 6.1e-06;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHOIQYFGQF 18
| | | | | | | | | | | | | | | | | |
Db 1 AALEAKICHOIEYFGDF 18

RESULT 12

US-09-836-073-4

Sequence 4, Application US/09836073

Patent No. US20020173475A1

GENERAL INFORMATION:

APPLICANT: Dasgupta, Asim

APPLICANT: Das, S.

APPLICANT: Baidya, Narayan

TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

FILE REFERENCE: 220002054822

CURRENT APPLICATION NUMBER: US/09/836,073

CURRENT FILING DATE: 2002-10-24

PRIOR APPLICATION NUMBER: 09/316,630

PRIOR FILING DATE: 1999-05-21

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4

LENGTH: 18

TYPE: PRT

ORGANISM: Homo Sapiens

US-09-836-073-4

Query Match 76.5%; Score 75; DB 9; Length 18;
Best Local Similarity 80.0%; Pred. No. 6.1e-06;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 QAKICHOIQYFGQF 18
| | | | | | | | | | | | | | | | | |
Db 4 EAKICHOIEYFGDF 18

RESULT 13

US-09-836-073-7

Sequence 7, Application US/09836073

Patent No. US20020173475A1

GENERAL INFORMATION:

APPLICANT: Dasgupta, Asim

APPLICANT: Das, S.

APPLICANT: Baidya, Narayan

TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

FILE REFERENCE: 220002054822

CURRENT APPLICATION NUMBER: US/09/836,073

CURRENT FILING DATE: 2002-10-24

PRIOR APPLICATION NUMBER: 09/316,630

PRIOR FILING DATE: 1999-05-21

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 18

TYPE: PRT

ORGANISM: Homo Sapiens

US-09-836-073-7

Query Match 73.5%; Score 72; DB 9; Length 18;
Best Local Similarity 81.2%; Pred. No. 1.9e-05;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHOIQYFG 16
| | | | | | | | | | | | | | | | | |
Db 1 AALEAKICHOIEYFG 16

RESULT 14

US-09-836-073-8

Sequence 8, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Baidya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 18
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-836-073-8

Query Match 72.4%; Score 71; DB 9; Length 18;
Best Local Similarity 72.2%; Pred. No. 2.7e-05;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 AALQAKICHOIQYFGQF 18
| | | | | | | | | | | | | | | | | |
Db 1 AALEAKICHOIEQFGDF 18

RESULT 15

US-09-836-073-15

Sequence 15, Application US/09836073

Patent No. US20020173475A1

GENERAL INFORMATION:

APPLICANT: Dasgupta, Asim

APPLICANT: Das, S.

APPLICANT: Baidya, Narayan

TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

FILE REFERENCE: 220002054822

CURRENT APPLICATION NUMBER: US/09/836,073

CURRENT FILING DATE: 2002-10-24

PRIOR APPLICATION NUMBER: 09/316,630

PRIOR FILING DATE: 1999-05-21

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 15

LENGTH: 18

TYPE: PRT

ORGANISM: Xenopus

US-09-836-073-15

Query Match 66.3%; Score 65; DB 9; Length 18;
Best Local Similarity 68.8%; Pred. No. 0.00026;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 LQAKICHOIQYFGQF 18
| | | | | | | | | | | | | | | | | |
Db 3 LDTKICEQIEYFGDF 18

Search completed: April 23, 2003, 13:38:19
Job time : 10.3146 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:24 ; Search time 11.2247 Seconds
(without alignments)
47.183 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98

Sequence: 1 AALQAKICHQIYYFGQF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	88.8	18	4	US-09-316-630-3
2	87	88.8	18	4	US-09-316-630-4
3	59	60.2	38	4	US-08-974-549A-214
4	50	51.0	38	3	US-08-851-843A-25
5	50	51.0	38	4	US-08-974-549A-215
6	50	51.0	38	4	US-08-854-050-25
7	50	51.0	38	4	US-09-430-323-25
8	49.5	50.5	37	3	US-08-851-843A-24
9	49.5	50.5	37	4	US-08-854-050-24
10	49.5	50.5	37	4	US-09-430-323-24
11	42	42.9	1220	2	US-08-680-326-38
12	41	41.8	292	4	US-09-134-001C-3515
13	40	40.8	111	4	US-09-257-179-61
14	40	40.8	616	4	US-09-298-367B-11
15	39	39.8	39	3	US-08-851-843A-26
16	39	39.8	39	4	US-08-974-549A-216
17	39	39.8	39	4	US-08-854-050-26
18	39	39.8	39	4	US-09-430-323-26
19	38.5	39.3	740	4	US-09-323-872A-23
20	38.5	39.3	864	4	US-09-323-872A-28
21	38	38.8	305	1	US-08-420-235B-25
22	38	38.8	305	4	US-08-793-624-25
23	38	38.8	305	5	PCT-US95-10194-25
24	38	38.8	984	4	US-09-287-354-2
25	38	38.8	1189	4	US-09-287-354-3
26	38	38.8	1189	4	US-09-287-354-4
27	38	38.8	1207	4	US-09-287-354-5

```

28 37 37.8 40 1 US-08-168-091A-43 Sequence 43, Appl
29 37 37.8 79 2 US-07-885-089B-31 Sequence 31, Appl
30 37 37.8 87 1 US-07-847-743B-18 Sequence 18, Appl
31 37 37.8 87 1 US-08-456-201-18 Sequence 18, Appl
32 37 37.8 87 2 US-08-330-161-16 Sequence 16, Appl
33 37 37.8 87 2 US-08-456-241-18 Sequence 18, Appl
34 37 37.8 87 2 US-08-440-401-16 Sequence 16, Appl
35 37 37.8 87 2 US-08-419-878B-16 Sequence 16, Appl
36 37 37.8 87 4 US-09-173-480-16 Sequence 16, Appl
37 37 37.8 87 5 PCT-US92-04295A-18 Sequence 18, Appl
38 37 37.8 147 6 5202428-8 Patent No. 5202428
39 37 37.8 214 4 US-08-861-774E-66 Sequence 66, Appl
40 37 37.8 243 2 US-07-885-089B-4 Sequence 4, Appl
41 37 37.8 243 2 US-07-885-089B-9 Sequence 9, Appl
42 37 37.8 243 6 5202428-2 Patent No. 5202428
43 37 37.8 543 4 US-09-535-008-63 Sequence 63, Appl
44 37 37.8 577 4 US-09-535-008-61 Sequence 61, Appl
45 37 37.8 584 4 US-09-102-528-12 Sequence 12, Appl

```

ALIGNMENTS

```

RESULT 1
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316.630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3
Query Match 88.8%; Score 87; DB 4; Length 18;
Best Local Similarity 83.3%; Pred. No. 4.9e+08;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 AALQAKICHQIYYFGQF 18
| | | | | | | | | | | | | | | |
DB 1 AALEAKICHQIYYFGDF 18
| | | | | | | | | | | | | | | |
RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316.630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

```

;; PRIOR FILING DATE: 1997-10-06
;; PRIOR APPLICATION NUMBER: 08/321,427
;; PRIOR FILING DATE: 1994-10-11
;; PRIOR APPLICATION NUMBER: 60/086,527
;; PRIOR FILING DATE: 1998-05-22
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
;; OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 88.8%; Score 87; DB 4; Length 18;
Best Local Similarity 83.3%; Pred. No. 4.9e-08;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHOIQIYFGQF 18
|||:|||||:|||||
Db 1 AALEAKICHOIYFYGDF 18

RESULT 3
US-08-974-549A-214
; Sequence 214, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997

;; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 214:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-549A-214

Query Match 60.2%; Score 59; DB 4; Length 38;
Best Local Similarity 75.0%; Pred. No. 0.0039;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIYIYFGQF 18
||||:|||||
Db 1 ICHQXEYFYGDF 12

RESULT 4
US-08-851-843A-25
; Sequence 25, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419

;; FILING DATE: 18-APR-1997
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002930US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
US-08-851-843A-25

Query Match 51.0%; Score 50; DB 3; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.11;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICHQIQYFG 16
||| ||| |||
Db 1 ICEQIEYFG 10

RESULT 5
US-08-974-549A-215
; Sequence 215, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/854,050
;; FILING DATE: 09-MAY-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/911,312
;; FILING DATE: 14-AUG-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/912,951
;; FILING DATE: 14-AUG-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/915,503
;; FILING DATE: 14-AUG-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US97/17618
;; FILING DATE: 01-OCT-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US97/17885
;; FILING DATE: 01-OCT-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph Ted
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002610US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0300
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 215:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-974-549A-215

Query Match 51.0%; Score 50; DB 4; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.11;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICHQIQYFG 16
||| ||| |||
Db 1 ICEQIEYFG 10

RESULT 6
US-08-854-050-25
; Sequence 25, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/851,843
;; FILING DATE: 06-MAY-1997
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/846,017
;; FILING DATE: 25-APR-1997
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/844,419
;; FILING DATE: 18-APR-1997
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/844,419
;; FILING DATE: 18-APR-1997
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002930US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; US-08-854-050-25

Query Match 51.0%; Score 50; DB 4; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.11;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIQYVFG 16
|||:|||||
Db 1 ICEQIEYVFG 10

RESULT 7
US-09-430-323-25
; Sequence 25, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
; TITLE OF INVENTION: NO. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050

;; FILING DATE: 09-MAY-1997
;; APPLICATION NUMBER: US 08/851,843
;; FILING DATE: 06-MAY-1997
;; APPLICATION NUMBER: US 08/846,017
;; FILING DATE: 25-APR-1997
;; APPLICATION NUMBER: US 08/844,419
;; FILING DATE: 18-APR-1997
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002930US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match 51.0%; Score 50; DB 4; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.11;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIQYVFG 16
|||:|||||
Db 1 ICEQIEYVFG 10

RESULT 8
US-08-851-843A-24
; Sequence 24, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
; TITLE OF INVENTION: NO. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-0029300S
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 37 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; US-08-851-843A-24

Query Match 50.5%; Score 49.5; DB 3; Length 37;
Best Local Similarity 75.0%; Pred. No. 0.13;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIQYFQF 18
DB 1 ICHQ-EYFGDF 11

RESULT 9
US-08-854-050-24
; Sequence 24, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:

;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-0029300S
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 37 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; US-08-854-050-24

Query Match 50.5%; Score 49.5; DB 4; Length 37;
Best Local Similarity 75.0%; Pred. No. 0.13;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIQYFQF 18
DB 1 ICHQ-EYFGDF 11

RESULT 10
US-09-430-323-24
; Sequence 24, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-430-323-24

Query Match 50.5%; Score 49.5; DB 4; Length 37;
Best Local Similarity 75.0%; Pred. No. 0.13;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIQYFGQF 18
||| :||| |
Db 1 ICHQ-EYFGDF 11

RESULT 11

US-08-680-326-38
; Sequence 38, Application US/08680326
; Patent No. 5925733
; GENERAL INFORMATION:
; APPLICANT: ROSE, TIMOTHY M.
; APPLICANT: BOSCH, MARNIX
; APPLICANT: STRAND, KURT
; APPLICANT: TODARO, GEORGE J.
; TITLE OF INVENTION: DNA POLYMERASE OF GAMMA HERPES VIRUSES
; TITLE OF INVENTION: ASSOCIATED WITH KAPOSI'S SARCOMA AND RETROPERITONEAL
; TITLE OF INVENTION: FIBROMATOSIS
; NUMBER OF SEQUENCES: 152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680,326
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20001.00
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141

INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1220 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-680-326-38

Query Match 42.9%; Score 42; DB 2; Length 1220;
Best Local Similarity 60.0%; Pred. No. 81;
Matches 6; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 7 ICHQIQYFG 16
||| :||| |
Db 248 VCHTLYYFG 257

RESULT 12

US-09-134-001C-3515
; Sequence 3515, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3515
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3515

Query Match 41.8%; Score 41; DB 4; Length 292;
Best Local Similarity 61.5%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 3; Indels 3; Gaps 0;

QY 6 KICHQIQYFGQF 18
||| :||| |
Db 27 KIENQIQYLEQF 39

RESULT 13

US-09-257-179-61
; Sequence 61, Application US/09257179
; Patent No. 6410709
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 29 Human Secreted Proteins
; FILE REFERENCE: P2015P1
; CURRENT APPLICATION NUMBER: US/09/257,179
; CURRENT FILING DATE: 1999-02-25
; EARLIER APPLICATION NUMBER: PCT/US98/17709
; EARLIER FILING DATE: 1998-08-27
; EARLIER APPLICATION NUMBER: 60/056,270
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,271
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,247
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,073
; EARLIER FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (111)
; OTHER INFORMATION: Xaa equals stop translation
US-09-257-179-61

Query Match 40.8%; Score 40; DB 4; Length 111;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 9 HQIQYFGQF 18
||| :||| |
Db 86 HLPDYFGQF 95

```

RESULT 14
US-09-298-367B-11
; Sequence 11, Application US/09298367B
; Patent No. 6180112
; GENERAL INFORMATION:
; APPLICANT: Highlander, Sarah K.
; APPLICANT: Federova, Natalie D.
; TITLE OF INVENTION: PASTEURILLA HAEMOLYTICA VACCINE
; FILE REFERENCE: BCM-03728
; CURRENT APPLICATION NUMBER: US/09/298,367B
; CURRENT FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 08/834,455
; PRIOR FILING DATE: 1997-04-15
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 11
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Pasteurella haemolytica
US-09-298-367B-11

```

```

Query Match 40.8%; Score 40; DB 4; Length 616;
Best Local Similarity 50.0%; Pred. No. 85;
Matches 9; Conservative 3; Mismatches 4; Indels 2; Gaps 1;
QY 3 LOAK--ICHQIQYVFG 18
||||| : | :| :|
Db 262 LQAKDILGHVYVFLGQF 279

```

```

RESULT 15
US-08-851-843A-26
; Sequence 26, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-851-843A-26
Query Match 39.8%; Score 39; DB 3; Length 39;
Best Local Similarity 60.0%; Pred. No. 7.2;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 7 ICHQIQYVFG 16
| | | | |
Db 1 ILRQVEYVFG 10
Search completed: April 23, 2003, 13:36:31
Job time : 12.2247 secs

```

GenCore version 5.1.4.p5_4578
Copyright. (c) 1993 - 2003. Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:23 ; Search time 28.5169 Seconds
(without alignments)
84.109 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98

Sequence: 1 AALQAKICHIQYVFGQF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	87	88.8	18	21	AA52200 Human la autoantigen
2	87	88.8	92	21	AAW01351 Human secreted pro
3	87	88.8	408	17	AAW03716 Human autoantigen
4	87	88.8	439	22	ABG08417 Novel human diagen
5	87	88.8	460	21	AAW58987 Breast and ovarian
6	87	88.8	460	23	ABP41511 Human ovarian anti
7	64	65.3	21	14	AAW43394 La/SSB epitope 17.
8	47	48.0	609	20	AAW87995 An alternatively s
9	47	48.0	609	21	AAW70961 Human Ras signalli
10	47	48.0	728	20	AAW87994 A human MCG7 prote

11	46	46.9	608	21	AAW70960
12	42	42.9	175	18	AAW20649
13	42	42.9	324	19	AAW98640
14	42	42.9	324	19	AAW71525
15	42	42.9	542	22	AAW46301
16	42	42.9	551	23	ABP41995
17	42	42.9	557	22	AAW46354
18	42	42.9	2410	18	AAW19723
19	42	42.9	2480	18	AAW19724
20	42	42.9	2644	18	AAW13152
21	42	42.9	2644	20	AAW84271
22	41	41.8	120	22	AAW74574
23	41	41.8	292	23	ABP38670
24	41	41.8	296	22	ABW60055
25	41	41.8	544	22	AAW75090
26	41	41.8	569	21	AAW57005
27	41	41.8	628	22	ABW64933
28	41	41.8	697	22	AAW64616
29	40	40.8	111	20	AAW04314
30	40	40.8	423	21	AAW27317
31	40	40.8	425	21	AAW27316
32	40	40.8	477	21	AAW49578
33	40	40.8	479	21	AAW49577
34	40	40.8	616	19	AAW83023
35	40	40.8	2169	22	AAW91701
36	39	39.8	40	22	ABW95660
37	39	39.8	40	22	AAW94199
38	39	39.8	88	22	AAW40039
39	39	39.8	88	23	ABW06216
40	39	39.8	111	22	AAU19464
41	39	39.8	111	23	ABP51350
42	39	39.8	188	21	AAW58237
43	39	39.8	188	22	AAW03032
44	39	39.8	191	22	AAW41825
45	39	39.8	232	21	AAW58236

ALIGNMENTS

RESULT 1
AA52200
ID AA52200 standard; peptide; 18 AA.

XX AC AA52200;

XX DT. 14-MAR-2000 (first entry)

XX DE Human la autoantigen peptide (LAP).

XX KW La autoantigen; LAP; internal ribosome entry site; IRES; translation;
KW viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
KW coronavirus; hepatitis virus; rhadovirus; adenovirus; coxsackie virus;
KW parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
KW vesicular stomatitis virus.

XX OS Homo sapiens.

XX PN WO9961613-A2.

XX PD 02-DEC-1999.

XX PF 21-MAY-1999; 99WO-US11281.

XX PR 22-MAY-1998; 98US-0086527.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Das S, Dasgupta A;

XX DR WPI; 2000-062712/05.

XX XX

PT New yeast inhibitory peptide useful for inhibiting viral protein
PT translation and replication -
XX
PS Claim 5; Page 57; 81pp; English.
XX
CC This sequence is the La autoantigen binding domain (LAP). LAP is a
CC cellular protein which binds to an internal ribosome entry site (IRES).
CC The peptide is used to inhibit mRNA translation, viral protein
CC translation or viral replication. Viral mRNA translation is initiated at
CC the IRES and the LAP peptide prevents translation initiation factors
CC from binding at the site. The peptide can be used alone or in combination
CC with an inhibitor RNA (IRNA see A245200). The LAP peptide is useful as
CC an antiviral agent, which works through the inhibition of mRNA
CC translation, especially viral mRNA. Examples of viruses which can be
CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
CC viruses, rhinovirus, adenovirus, and parainfluenza virus, poliovirus,
CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
CC The peptide also inhibits replication of the above viruses. The LAP
CC peptide selectively inhibits viral protein translation, and is therefore
CC not toxic to the host cell.
XX
SQ Sequence 18 AA;

Query Match 88.8%; Score 87; DB 21; Length 18;
Best Local Similarity 83.3%; Pred. No. 3.1e-07;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
| | | | | | | | | | | | | | | | | |
Db 1 AALEAKICHQIEYFGDF 18

RESULT 2
AAG01351
ID ARG01351 standard; Protein; 92 AA.
XX
AC AAG01351;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein, SEQ ID NO: 5432.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.
XX
XX Homo sapiens.
XX
XX EPI033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
XX N-PSDB; AAC01357.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 13; SEQ ID 5432; 71pp + CD-ROM; English.
XX
XX The present sequence is a polypeptide encoded by one of a large number
XX of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
XX were prepared from total human RNAs or polyA+ RNAs derived from 30

CC different tissues. EST sequences usually correspond mainly to the 3',
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5',
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5',
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.
XX
SQ Sequence 92 AA;

Query Match 88.8%; Score 87; DB 21; Length 92;
Best Local Similarity 83.3%; Pred. No. 1.7e-06;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
| | | | | | | | | | | | | | | | | |
Db 11 AALEAKICHQIEYFGDF 28

RESULT 3
AAW03716
ID AAW03716 standard; protein; 408 AA.
XX
XX AAW03716;
XX
XX 12-MAR-1997 (first entry)
XX
XX Human autoantigen La(SS-B).
XX
XX Autoimmune disease; La autoantigen; Sjogren's syndrome;
KW systemic lupus erythematosus; diagnosis.
XX
XX Homo sapiens.
XX
XX US5541291-A.
XX
XX 30-JUL-1996.
XX
XX 31-DEC-1984; 84US-0687908.
XX
XX 27-MAY-1987; 87US-0054871.
XX
XX 31-DEC-1984; 84US-0687908.
XX
XX (UYDU-) UNIV DUKE.
XX
XX Keene JD;
XX
XX WPI; 1996-362015/36.
XX
XX Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
XX overlap syndrome - useful for diagnosis and treatment of autoimmune
XX diseases
XX
XX Disclosure; Columns 15-16; 21pp; English.
XX
XX The human lupus antigen (La) is diagnostic for Sjogren's syndrome,
XX as well as occurring in systemic lupus erythematosus patients. The
XX La protein is clinically related to the Ro protein that is highly
XX common among autoimmune patients. La and Ro antigens sometimes
XX reside on the same cellular ribonucleoprotein particle; most La
XX patients contain some Ro antibodies and vice versa. La cDNA has
XX been isolated from a human liver library.
XX
XX Sequence 408 AA;

Query Match 88.8%; Score 87; DB 17; Length 408;
Best Local Similarity 83.3%; Pred. No. 7.8e-06;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18

Db 11 AALEAKICHQIEYFGDF 28
|||||:|||||:|||||

RESULT 4
ABG08417
ID ABG08417 standard; Protein; 439 AA.

AC ABG08417;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #8408.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Dmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS72604.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX PS Claim 20; SEQ ID NO 38776; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 439 AA;

Query Match 88.8%; Score 87; DB 22; Length 439;

Best Local Similarity 83.3%; Pred. NO. 8.4e-06;

Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 AALQAKICHQIEYFGDF 18
|||||:|||||:|||||

Db 41 AALEAKICHQIEYFGDF 58

RESULT 5

AAB58987

XX ID AAB58987 standard; Protein; 460 AA.

XX AC AAB58987;

XX DT 27-MAR-2001 (first entry)

XX DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 695.

XX KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neutropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.

XX OS Homo sapiens.

XX PN WO2000055173-A1.

XX PD 21-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US05881.

XX PR 12-MAR-1999; 99US-0124270.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX DR WPI; 2000-611515/58.

XX DR N-PSDB; AAF21890.

XX PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases

XX PS Claim 11; Page 1149-11150; 1299pp; English.

XX CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC neutropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiant activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.

XX SQ Sequence 460 AA;

Query Match 88.8%; Score 87; DB 21; Length 460;

Best Local Similarity 83.3%; Pred. NO. 8.8e-06;

Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 AALQAKICHQIEYFGDF 18
|||||:|||||:|||||

CC autoantibody and may therefore be used as vaccines.

XX Sequence 21 AA;

Query Match 65.3%; Score 64; DB 14; Length 21;
Best Local Similarity 83.3%; Pred. No. 0.0018;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 7 ICHQIQYFGQF 18
DB 1 ICHQIEYFGDF 12

RESULT 8

AAW87995
ID AAW87995 standard; Protein: 609 AA.

XX AC AAW87995;

XX DT 15-APR-1999 (first entry)

XX DE An alternatively spliced human MCG7 protein.

XX KW MCG4 protein; gene regulatory function; heat shock protein;
XX KW guanine nucleotide exchange factor protein; MCG7 protein;
XX KW heat shock-binding protein; MCG18 protein; zinc finger protein;
XX KW cancer.

XX OS Homo sapiens.

XX PN W09853061-A1.

XX PD 26-NOV-1998.

XX PF 22-MAY-1998; 98WO-AU00380.

XX PR 22-JAN-1998; 98AU-0001460.

XX PR 23-MAY-1997; 97AU-0006972.

XX PR 23-MAY-1997; 97AU-0006973.

XX PR 23-MAY-1997; 97AU-0006974.

XX PR 22-JAN-1998; 98AU-0001458.

XX PR 22-JAN-1998; 98AU-0001459.

XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

XX PI Gartside M, Grimmond S, Hancock J, Hayward N, Silins G;

XX DR WPI: 1999-070146/06.

XX DR N-PSDB; AAX04553.

XX PT New gene-expression regulatory genes, mcg4, mcg7, and mcg18 - encode
XX PT a zinc finger protein, a GEF, and a heat shock or heat shock binding
XX PT protein, useful to detect and treat cancer
XX PS Claim 5; Fig 13b; 80pp; English.
XX CC The present sequence represents a MCG7 protein. The protein has gene
XX CC regulatory functions, and has homology to a heat shock protein or
XX CC heat shock-binding protein. The specification also describes MCG4,
XX CC which is homologous to guanine nucleotide exchange factor protein,
XX CC and MCG18, which is homologous to a zinc finger protein.
XX CC Detection of mutations in the MCG genes can be used to identify the
XX CC propensity for various types of cancer, and to treat, arrest, or
XX CC otherwise ameliorate, the effects of a cancer in an animal or bird.

XX Sequence 609 AA;

Query Match 48.0%; Score 47; DB 20; Length 609;

Best Local Similarity 41.2%; Pred. No. 31;

Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 2 ALQAKICHOIQYFGQF 18

||| ||| ::|

DB 70 SLOVKTCHLVRYWISAF 86

RESULT 9

AAV70961

XX ID AAV70961 standard; Protein: 609 AA.

XX AC AAV70961;

XX DT 09-AUG-2000 (first entry)

XX DE Human Ras signalling pathway associated protein CalDAG-GEFI.

XX KW Human; Ras signalling pathway; CalDAG-GEFI; calcium; DAG; diacylglycerol;
XX KW GEF; guanine nucleotide exchange factor; Rap1A; diagnosis; treatment;
XX KW CalDAG-GEF-associated disorder; drug; transgenic animal model;
XX KW Ras-associated cancer; protein therapy.

XX OS Homo sapiens.

XX FH Key. Location/Qualifiers

XX FT 150..173

XX FT /label= SCR1

XX FT /note= "Structurally conserved region which is

XX FT highly homologous to Ras-superfamily GEFs"

XX FT 220..262

XX FT /label= SCR2

XX FT /note= "Structurally conserved region which is

XX FT highly homologous to Ras-superfamily GEFs"

XX FT 299..320

XX FT /label= SCR3

XX FT /note= "Structurally conserved region which is

XX FT highly homologous to Ras-superfamily GEFs"

XX FT 433..452

XX FT /label= EF-hand

XX FT /note= "calcium binding motif"

XX FT 499..548

XX FT /label= DAG/phorbol_ester-binding_domain

XX FT /note= "present in most protein kinase C family members"

XX PN W0200024768-A2.

XX PD 04-MAY-2000.

XX PF 22-OCT-1999; 99WO-US24826.

XX PR 23-OCT-1998; 98US-0105507.

XX PR 16-NOV-1998; 98US-0108685.

XX PA (MASI) MASSACHUSETTS INST TECHNOLOGY.

XX PI Kawasaki H, Graybiel A, Housman D;

XX DR WPI: 2000-350690/30.

XX DR N-PSDB; AAD00311.

XX PT Isolated nucleic acid comprises nucleotide sequence encoding protein
XX PT selected from normal or mutant CalDAG-guanine nucleotide exchange
XX PT factor

XX PS Claim 64; Page 90-92; 128pp; English.

XX CC The present sequence is a CalDAG-guanine nucleotide exchange factor I
XX CC (GEFI) from a human frontal cortex and U937 lambdaZAPII cDNA libraries.
XX CC The CalDAG-GEFI has substrate specificity for

XX CC Rap1A, dual binding domains for calcium and diacylglycerol (DAG) and
XX CC highly expressed in brain, particularly in brain basal ganglia pathways
XX CC and their axon-terminal regions. Expression of CalDAG-GEFI activates
XX CC Rap1A and inhibits Ras-dependent activation of the extracellular-signal
XX CC regulated kinase/mitogen-activated protein (ERK/MAP) kinase cascade in
XX CC 293T cells. The CalDAG-GEF proteins play an important role in determining
XX CC the relative activation of Ras and Rap1 signalling induced by calcium and
XX CC DAG mobilisation in brain and haematopoietic organs. The present

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:34:48 ; Search time 20.0225 Seconds
(without alignments)
185.234 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98
Sequence: 1 AALQAKICHQIQYYFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 9297

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	31.6	24	5 Q9BM09	Q9bm09 spongilla 1
2	28.5	29.1	23	5 Q9S595	Q9S595 drosophila
3	28	28.6	24	2 Q05616	Q05616 staphylococ
4	27	27.6	13	5 Q917F8	Q917F8 drosophila
5	27	27.6	24	6 Q28430	Q28430 gorilla gor
6	27	27.6	24	6 Q28781	Q28781 pongo pygma
7	25	25.5	16	2 Q9E963	Q9E963 helicobacte
8	25	25.5	18	4 Q16028	Q16028 homo sapien
9	25	25.5	19	6 Q9EDD6	Q9edd6 ovis aries
10	25	25.5	23	13 Q9PRV2	Q9prv2 gallus gall
11	25	25.5	25	13 Q9PRV1	Q9prv1 gallus gall
12	24.5	25.0	24	9 Q38270	Q38270 bacterioph
13	24	24.5	12	10 Q9M433	Q9m433 lotus japon
14	24	24.5	17	11 P97758	P97758 mus musculu
15	24	24.5	19	2 Q9R4B9	Q9r4b9 streptococ
16	24	24.5	19	8 Q9XMB5	Q9xmb5 aegilops ta

17	24	24.5	19	12 Q83273	Q83273 cucumber mo
18	24	24.5	24	7 Q9TNS7	Q9tns7 homo sapien
19	24	24.5	24	7 Q9TNS6	Q9tns6 homo sapien
20	24	24.5	25	5 Q9TNC6	Q9txc6 drosophila
21	24	24.5	25	8 Q9TGB8	Q9tgb8 alnus crisp
22	24	24.5	25	8 Q9TGB7	Q9tgb7 alnus glut
23	24	24.5	25	8 Q9TGB6	Q9tgb6 alnus marit
24	24	24.5	25	8 Q9TGB5	Q9tgb5 betula alle
25	24	24.5	25	8 Q9TGB4	Q9tgb4 betula gian
26	24	24.5	25	8 Q9TGB3	Q9tgb3 betula papy
27	24	24.5	25	8 Q9TGB2	Q9tgb2 betula verr
28	24	24.5	25	8 Q9TGB1	Q9tgb1 betula pube
29	24	24.5	25	8 Q9TGB0	Q9tgb0 corylus ave
30	24	24.5	25	8 Q9TGA9	Q9tga9 corylus col
31	24	24.5	25	8 Q9TGA8	Q9tga8 corylus cor
32	24	24.5	25	8 Q9TGA7	Q9tga7 ostrya virg
33	24	24.5	25	8 Q9TGA6	Q9tga6 quercus rub
34	24	24.5	25	8 Q9TGA3	Q9tga3 carpinus ca
35	24	24.5	25	13 Q90Z49	Q90z49 haplochromi
36	23	23.5	8	4 Q9BY55	Q9byy5 homo sapien
37	23	23.5	8	6 Q9BFA0	Q9bfa0 macaca mula
38	23	23.5	8	6 Q9BF99	Q9bf99 hylobates c
39	23	23.5	9	12 Q9E1U7	Q9elu7 hepatitis b
40	23	23.5	13	12 Q9E1V4	Q9elv4 hepatitis b
41	23	23.5	13	12 Q9E1V3	Q9elv3 hepatitis b
42	23	23.5	13	12 Q9E1V2	Q9elv2 hepatitis b
43	23	23.5	13	12 Q9E1V1	Q9elv1 hepatitis b
44	23	23.5	13	12 Q9E1V0	Q9elv0 hepatitis b
45	23	23.5	13	12 Q9E1U9	Q9elu9 hepatitis b

ALIGNMENTS

RESULT 1

Q9BM09	PRELIMINARY;	PRT;	24 AA.
ID Q9BM09			
AC Q9BM09			
DT 01-JUN-2001 (TREMBlrel. 17, Created)			
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)			
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)			
DE Gypsy-like reverse transcriptase (Fragment).			
OS Spongilla lacustris (Freshwater sponge).			
OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;			
OC Haplosclerida; Spongillidae; Spongilla.			
OX NCBI_TaxID=6055;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC TRANSPOSON-GRT-G7 RETROTRANSPOSON;			
RX MEDLINE-20570504; PubMed-11121049;			
RA Arkhipova I., Meselson M.;			
RT "Transposable elements in sexual and ancient asexual taxa.";			
RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).			
DR EMBL; AR013997; AAG59969.1; -.			
KW RNA-directed DNA polymerase.			
FT NON_TER 1			
FT NON_TER 24			
SQ SEQUENCE 24 AA; 2754 MW; B588512EDE7A251A CRC64;			

Query Match 31.6%; Score 31; DB 5; Length 24;
Best Local Similarity 30.8%; Pred. No. 2.3e+02;
Matches 4; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 LQAKICHQIQYYF 15

DB 11 LMDTVCHGLEFVF 23

RESULT 2

Q9S595	PRELIMINARY;	PRT;	23 AA.
ID Q9S595			
AC Q9S595			
DT 01-DEC-2001 (TREMBlrel. 19, Created)			


```

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE GM05003P
GN ANTP OR BG:DS07700.1 OR CG1028.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Ceiniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY060900; AAL28448.1; -
DR FLYBase; FBgn0000095; Antp.
SQ SEQUENCE 23 AA; 2820 MW; 6F16D52A47F69139 CRC64;

Query Match 29.1%; Score 28.5; DB 5; Length 23;
Best Local Similarity 53.8%; Pred. No. 6e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 3 LQAKI-CHQIQYF 14
Db 11 IQIKHPHIVQY 23

RESULT 3
Q05616 PRELIMINARY; PRT; 24 AA.
AC Q05616;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE 3-dehydroquinase synthase (EC 4.6.1.3) (Fragment).
GN AKOB.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=8325-4;
RX MEDLINE=93381456; PubMed=8371108;
RA O'Connell C.M., Pattee P., Foster T.J.;
RT "Sequence and mapping of the araA gene of Staphylococcus aureus 8325-4.";
RL J. Gen. Microbiol. 139:1449-1460(1993).
CC -1- CATALYTIC ACTIVITY: 7-PHOSPHO-3-DEOXY-ARABINO-HEPTULOSONATE - 3-DEHYDROQUINATE + ORTHOPHOSPHATE.
CC -1- PATHWAY: SECOND STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
DR EMBL; L05004; AAA71896.1; -
KW Aromatic amino acid biosynthesis; Lyase.
FT NON_TER 1
SQ SEQUENCE 24 AA; 2857 MW; 1C4C8D20F7A4856E CRC64;

Query Match 28.6%; Score 28; DB 2; Length 24;
Best Local Similarity 50.0%; Pred. No. 7.6e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 CHQIQYF 15
Db 16 CEQLKTYF 23

RESULT 4
Q017F8 PRELIMINARY; PRT; 13 AA.
ID Q017F8
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

```

```

AC Q017F8;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
GN CG18278 protein (Fragment).
DE CG18278.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003818; AAG2274.1; -
DR FLYBase; FBgn0033836; CG18278.
FT NON_TER 1
SQ SEQUENCE 13 AA; 1554 MW; 5D52855A93735EB7 CRC64;

Query Match 27.6%; Score 27; DB 5; Length 13;
Best Local Similarity 36.4%; Pred. No. 6.3e+02;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 LQAKI-CHQIQY 13
Db 2 LQAGLCYRTNF 12

RESULT 5
Q28430 PRELIMINARY; PRT; 24 AA.
ID Q28430
AC Q28430;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

```

DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Relaxin (Fragment).
 GN RLX.
 OS Gorilla gorilla (gorilla).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 OX NCBI_TaxID=9593;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RG YK;
 RA Evans B.B.A., Fu P., Tregear G.G.W.;
 RL "Characterisation of primate relaxin genes.";
 RT J. Mol. Endocrinol. 0:0-0(1993).
 CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 DR EMBL; 227228; CA81742.1; -;
 DR HSSP; P04090; 6RLX.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 FT NON_TER 1
 FT 1
 SQ SEQUENCE 24 AA; 2673 MW; 5615D4B92C1DFD24 CRC64;
 Query Match 27.6%; Score 27; DB 6; Length 24;
 Best Local Similarity 55.6%; Pred. No. 1.1e+03;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 AALQAKICH 9
 Db :||| |
 4 SALANKCCH 12
 RESULT 6
 Q28781 PRELIMINARY; PRT; 24 AA.
 AC Q28781;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Relaxin (Fragment).
 GN RLX.
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RG YK 287;
 RA Evans B.B.A., Fu P., Tregear G.G.W.;
 RL "Characterisation of primate relaxin genes.";
 RT J. Mol. Endocrinol. 0:0-0(1993).
 CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 DR EMBL; 227229; CA81743.1; -;
 DR HSSP; P04090; 6RLX.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 FT NON_TER 1
 FT 1
 SQ SEQUENCE 24 AA; 2673 MW; 5615D4B92C1DFD24 CRC64;
 Query Match 27.6%; Score 27; DB 6; Length 24;
 Best Local Similarity 55.6%; Pred. No. 1.1e+03;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 AALQAKICH 9
 Db :||| |
 4 SALANKCCH 12
 RESULT 7
 Q28781 PRELIMINARY; PRT; 16 AA.
 ID Q28781
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

AC Q98963;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
 DE FepC (Fragment).
 GN FEP.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=F31;
 RA MEDLINE=98453456; PubMed=9780260;
 RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,
 RA Kuriyama M.;
 RT "Full-length sequence analysis of the vacA gene from cytotoxic and
 RT noncytotoxic Helicobacter pylori.";
 RL J. Infect. Dis. 178:1391-1398(1998).
 DR EMBL; AF049623; AAD04263.1; -;
 FT NON_TER 1
 FT 1
 SQ SEQUENCE 16 AA; 1938 MW; 3C33D403AEE61428 CRC64;
 Query Match 25.5%; Score 25; DB 2; Length 16;
 Best Local Similarity 44.4%; Pred. No. 1.7e+03;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 10 QIQYYFGQF 18
 Db :||| |
 7 QVEYYAFNF 15
 RESULT 8
 Q16028 PRELIMINARY; PRT; 18 AA.
 AC Q16028;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
 DE OCRL-1 protein (Fragment).
 GN OCRL-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=93278398; PubMed=8504307;
 RA Leahey A.M., Charnas L.R., Nussbaum R.L.;
 RT "Nonsense mutations in the OCRL-1 gene in patients with the
 RT oculocerebrorenal syndrome of Lowe.";
 RL Hum. Mol. Genet. 2:461-463(1993).
 DR EMBL; S62084; AAD13933.1; -;
 FT NON_TER 1
 FT 1
 SQ SEQUENCE 18 AA; 2055 MW; 61356C825F7AF33D CRC64;
 Query Match 25.5%; Score 25; DB 4; Length 18;
 Best Local Similarity 36.4%; Pred. No. 1.9e+03;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AALQAKICHQI 11
 Db :||| |
 1 SAYDPRICRQL 11
 RESULT 9
 Q9BDD6 PRELIMINARY; PRT; 19 AA.
 ID Q9BDD6
 AC Q9BDD6;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

DE Choline/ethanolamine kinase isolog (Fragment).
GN CHOLINE KINASE.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Price N.T., Cameron J.M., van der Leij F.R., Zammit V.A.;
RT "Analysis of the promoter region for the ovine carnitine
RT palmitoyltransferase IB gene."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ288906; CAC36323.1; -;
DR EMBL; AJ288907; CAC36324.1; -;
KW Kinase.
FT NON_TER
SQ SEQUENCE 19 AA; 2328 MW; 938A7B7F1636318C CRC64;
Query Match 25.5%; Score 25; DB 6; Length 19;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 10 QIQYFQ 17
Db 1 RFQYFQ 8
:|||||
RESULT 10
Q9PRV2 PRELIMINARY; PRT; 23 AA.
AC Q9PRV2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Elastase 1 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE-95066525; PubMed-7976036;
RA Tusck F., Polanowski A., Guyonnet V., Long P.L., Travis J.;
RT "Affinity purification of chicken pancreas proteinases and their N-
RT terminal amino-acid sequences."
RL Acta Biochim. Pol. 41:174-177(1994).
DR HSP; P00772; LEAL.
SQ SEQUENCE 23 AA; 2491 MW; 441147DDB26A987C CRC64;
Query Match 25.5%; Score 25; DB 13; Length 23;
Best Local Similarity 66.7%; Pred. No. 2.4e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 11 IQYYFG 16
Db 18 LQYYS 23
:|||||
RESULT 11
Q9PRV1 PRELIMINARY; PRT; 25 AA.
AC Q9PRV1;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Elastase 2 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.

OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE-95066525; PubMed-7976036;
RA Tusck F., Polanowski A., Guyonnet V., Long P.L., Travis J.;
RT "Affinity purification of chicken pancreas proteinases and their N-
RT terminal amino-acid sequences."
RL Acta Biochim. Pol. 41:174-177(1994).
DR HSP; P00772; LEAL.
SQ SEQUENCE 25 AA; 2635 MW; BB72141147DDB26A CRC64;
Query Match 25.5%; Score 25; DB 13; Length 25;
Best Local Similarity 66.7%; Pred. No. 2.6e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 11 IQYYFG 16
Db 18 LQYYS 23
:|||||
RESULT 12
Q38270 PRELIMINARY; PRT; 24 AA.
AC Q38270;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Bacteriophage lambda mutant truncated regulatory protein (PyrI) gene
DE from E.coli, 3' end (Fragment).
OS Bacteriophage lambda.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10710;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-86143826; PubMed-3912513;
RA Cunin R., Jacobs A., Charlier D., Crabeel M., Herve G., Glansdorff N.,
RA Pierard A.;
RA "Structure-function relationship in allosteric aspartate
RT carboxyltransferase from Escherichia coli: 1. Primary structure of a
RT pyrI gene encoding a modified regulatory subunit."
RL J. Mol. Biol. 186:707-713(1985).
DR EMBL; M28579; AAA3252.1; -;
DR HSP; P00478; BAC.
FT NON_TER
SQ SEQUENCE 24 AA; 2905 MW; FD9349DF6F8159D0 CRC64;
Query Match 25.0%; Score 24.5; DB 9; Length 24;
Best Local Similarity 38.5%; Pred. No. 3e+03;
Matches 5; Conservative 4; Mismatches 1; Indels 1; Gaps 1;
QY 2 ALQAKICHQIYY 14
Db 8 ALRCKYCEK-EFY 19
:|||||
RESULT 13
Q9M433 PRELIMINARY; PRT; 12 AA.
AC Q9M433;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ENOD40-1 protein.
GN ENOD40-1.
OS Lotus japonicus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20427395; PubMed-10975655;

RA Fimetakis E., Kavroulakis N., Quaedvlieg N.E.M., Spaink H.P.,
RA Dinou M., Roussis A., Katinakis P.;
RT "Lotus japonicus contains two distinct ENOD40 genes that are expressed
in symbiotic, non-symbiotic and embryonic tissues.";
RL Mol. Plant Microbe Interact. 13:987-994(2000).
DR EMBL: AJ271787; CAB92978.1;
SQ SEQUENCE 12 AA; 1403 MW; 3C6955187CB046C3 CRC64;

Query Match 24.5%; Score 24; DB 10; Length 12;
Best Local Similarity 44.4%; Pred. No. 1.9e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 KICHQIQY 14
I:|:|:|:
Db 2 KLCWQISIH 10

RESULT 14
P97758 PRELIMINARY; PRT; 17 AA.
AC P97758;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Dm28 protein (Fragment).
GN H-2D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6BY-H-2BM28;
RX MEDLINE=97188477; PubMed=9037062;
RA Yun T.J., Melvold R.W., Pease L.R.;
RT "A complex major histocompatibility complex D locus variant generated
by an unusual recombination mechanism in mice.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:1384-1389(1997).
DR EMBL: U83304; AAB41021.1;
DR InterPro: IPR001039; MHC_I.
DR Pfam: PF00129; MHC_I; 1.
FT NON_TER 1
FT NON_TER 17
SQ SEQUENCE 17 AA; 1956 MW; F3149F377C16F196 CRC64;

Query Match 24.5%; Score 24; DB 11; Length 17;
Best Local Similarity 37.5%; Pred. No. 2.7e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 HOIQYFG 16
I:|:|:|:
Db 3 HTLQWYG 10

RESULT 15
Q9R4B9 PRELIMINARY; PRT; 19 AA.
AC Q9R4B9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Surface trypsin-resistant LADDERING protein (Fragment).
OS Streptococcus sp.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1306;
RN [1]
RP SEQUENCE.
RX MEDLINE=97047708; PubMed=8926097;
RA Lachnauer C.S., Madoff L.C.;
RT "A protective surface proteoly from type V group B streptococci shares
N-terminal sequence homology with the alpha C protein.";
RL Infect. Immun. 64:4255-4260(1996).

SQ SEQUENCE 19 AA; 2008 MW; F44CC89ED774EA3C CRC64;
Query Match 24.5%; Score 24; DB 2; Length 19;
Best Local Similarity 50.0%; Pred. No. 3e+03;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQ 12
I:|:|:|:|:
Db 7 ATLNTKTKNIQ 18

Search completed: April 23, 2003, 13:47:14
Job time : 21.0225 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:33:13 ; Search time 4.65169 Seconds
(without alignments)
160.495 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98

Sequence: 1 AALQAKICHQIYVFGQF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1520

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	24.5	11	1 CXL1_CONMR	P58807 conus marmo
2	24	24.5	12	1 TIN2_HOFTI	P82652 holostrac
3	24	24.5	13	1 CXL4_CONMR	P58810 conus marmo
4	24	24.5	18	1 MLB_SCYCA	P01206 scyllorhinu
5	24	24.5	23	1 XYCL_ACIGB	P46365 acinetobact
6	24	24.5	25	1 CXOB_CONMA	P05485 conus magus
7	24	24.5	25	1 SMBP_RAT	P80968 rattus norv
8	23	23.5	16	1 MLB_SQUAC	P01207 squallus aca
9	23	23.5	20	1 TL18_SPIOL	P82536 spinacia ol
10	23	23.5	25	1 ANDT_ANDAU	P56684 androctonus
11	22	22.4	8	1 AL17_CARMA	P81820 carcinus ma
12	22	22.4	9	1 AL11_CARMA	P81814 carcinus ma
13	22	22.4	14	1 LECB_PSOSC	P22584 psophocarpu
14	22	22.4	17	1 TPIS_PINPS	P81666 pinus pinas
15	22	22.4	21	1 BTX_ATRBI	P80163 atractaspis
16	22	22.4	21	1 STRD_ATREN	P13211 atractaspis
17	21	21.4	18	1 OBP_LYMDI	P34173 lymantria d
18	21	21.4	19	1 HBB2_UROHA	P18992 uromastyx h
19	21	21.4	23	1 PRO3_DACGL	P18690 dactylis gl
20	20	20.4	14	1 CXA1_CONCN	P56973 conus conso
21	20	20.4	15	1 PC20_BRANA	P81096 brassica na
22	20	20.4	16	1 CXA2_CONMA	P56636 conus magus
23	20	20.4	20	1 COG4_CHIOP	P34156 chionocete
24	20	20.4	21	1 NDK_CANAL	Q9ur66 candida alb
25	20	20.4	22	1 LP1_TRIWA	P24335 trimeresuru
26	20	20.4	22	1 LP2_TRIWA	P58930 trimeresuru
27	20	20.4	23	1 NUO5_SOLTU	P80262 solanum tub
28	20	20.4	25	1 NEUL_PIG	P34964 sus scrofa
29	19	19.4	10	1 GON2_CHEPR	P80678 chelyosoma
30	19	19.4	10	1 TKND_ONCMY	P28500 oncorhynch
31	19	19.4	11	1 TIN4_HOFTI	P82654 holostrac
32	19	19.4	11	1 TKNA_CHICK	P19850 gallus gall
33	19	19.4	11	1 TKNA_HORSE	P01290 equus cabal

34	19	19.4	12	1 N040_SESRO	O24369 sesbania ro
35	19	19.4	12	1 TIN3_HOFTI	P82653 holostrac
36	19	19.4	14	1 ADF_TENMO	P82965 tenebrio mo
37	19	19.4	14	1 MAST_VESLE	P01514 vespula lew
38	19	19.4	15	1 IRBP_CRISP	P12665 cricetidae
39	19	19.4	15	1 PCTS_PELAC	P80564 pelobacter
40	19	19.4	15	1 SPAN_HELAN	P81098 helianthus
41	19	19.4	20	1 LPTR_BACST	P05658 bacillus st
42	19	19.4	20	1 PNV2_PHONI	Q9twr5 phoneutria
43	19	19.4	21	1 TKNC_CARAU	P25421 carassius a
44	19	19.4	22	1 ATP6_COTJA	P50681 coturnix co
45	19	19.4	22	1 VGLG_RABVA	P15199 rabies viru

ALIGNMENTS

```
RESULT 1
CXL1_CONMR          STANDARD;          PRT;          11 AA.
ID   CXL1_CONMR
AC   P58807;
DT   15-JUN-2002 (Rel. 41, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Lambda-conotoxin CMrVIA.
OS   Conus marmoreus (Marble cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=42752;
RN   [1]
RP   SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC   TISSUE=Venom;
RX   MEDLINE=20564325; PubMed=10988292;
RA   Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA   Seow K.T., Bay B.-H.;
RT   "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT   pattern and protein folding. Isolation and characterization from the
RT   venom of Conus marmoreus."
RL   J. Biol. Chem. 275:39516-39522(2000).
CC   - FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC   - SUBCELLULAR LOCATION: Secreted.
CC   - TISSUE SPECIFICITY: Expressed by the venom duct.
CC   - MASS SPECTROMETRY: MW=1237.93; MW_ERR=0.21; METHOD=Electrospray.
CW   - SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KW   Neurotoxin; Toxin; Hydroxylation.
FT   DISULFID 2..11
FT   DISULFID 2..11
FT   DISULFID 3..8
FT   MOD_RES 10..10
FT   MOD_RES 10..10
SQ   SEQUENCE 11 AA; 1226 MW; 277AAC60B7232B58 CRC64;
      Query Match      24.5%; Score 24; DB 1; Length 11;
      Best Local Similarity 75.0%; Pred. No. 3.9e+02;
      Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY   6 KICH 9
     1:11
Db   6 KLCH 9

RESULT 2
TIN2_HOFTI          STANDARD;          PRT;          12 AA.
ID   TIN2_HOFTI
AC   P82652;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Tigerin-2.
OS   Holostrachus tigrinus (Indian bull frog) (Rana tigerina).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae;
OC   Holostrachus.
OX   NCBI_TaxID=103373;
RN   [1]
```

RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND DISULFIDE BONDS.

RC TISSUE-Skin;
 RX PubMed-11031261;
 RA Purna Sai K., Jagannadham M.V., Vairamani M., Raju N.P.,
 RA Devi A.S., Nagaraj R., Sitarum N.;
 RT "Tigerlins: novel antimicrobial peptides from the Indian frog Rana
 RT tigerina.";
 RL J. Biol. Chem. 276:2701-2707(2001).
 CC -1- FUNCTION: ANTIBACTERIAL ACTIVITY AGAINST B.SUBTILIS, E.COLI,
 CC S.AUREUS, M.LUTEUS, P.PUTIDA AND S.CEREVISIAE.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: SKIN.
 CC -1- MASS SPECTROMETRY: MW=1368; METHOD=MALDI.
 KW Amphibian skin; Antibiotic; Amidation.
 FT DISULFID 3 11
 FT MOD_RES 12 12
 FT MOD_RES 12 12
 FT MOD_RES 12 12
 SQ SEQUENCE 12 AA; 1369 MW; C2627D8EFD37605D CRC64;

Query Match 24.5%; Score 24; DB 1; Length 12;

Best Local Similarity 50.0%; Pred. No. 4.2e+02;

Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ALQAKICH 9

I: |||

Db 5 AIPLPICH 12

RESULT 3

ID CXL4_CONMR STANDARD; PRT; 13 AA.
 AC P58810;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lambda/chi-conotoxin MrIB (Chi-MrIB).
 OS Conus marmoreus (Marble cone).
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 CC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=42752;

RP SEQUENCE, SYNTHESIS, AND STRUCTURE BY NMR.

RX TISSUE=Venom;
 RX MEDLINE=21419681; PubMed=11528421;
 RA Sharpe I.A., Gehrmann J., Loughnan M.L., Thomas L., Adams D.A.,
 RA Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.,
 RA Lewis R.J.;
 RT "Two new classes of conopeptides inhibit the alpha1-adrenoceptor and
 RT noradrenaline transporter.";
 RL Nat. Neurosci. 4:902-907(2001).
 CC -1- FUNCTION: Inhibits the neuronal noradrenaline transporter.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -1- PTM: Exists in two forms, due to cis-trans isomerization at His-
 CC 11-Hyp-12.
 CC -1- MASS SPECTROMETRY: MW=1393.52; METHOD=Electrospray.
 CC -1- SIMILARITY: BELONGS TO THE CHI/LAMEDA-CONOTOXIN FAMILY.
 DR PDB; IIEO; 03-APR-02.
 KW Neurotoxin; Toxin; Hydroxylation; 3D-structure.
 FT DISULFID 4 13
 FT MOD_RES 5 10
 FT MOD_RES 12 12
 FT MOD_RES 12 12
 SQ SEQUENCE 13 AA; 1382 MW; 277AAC376EAD2B58 CRC64;

Query Match 24.5%; Score 24; DB 1; Length 13;

Best Local Similarity 75.0%; Pred. No. 4.6e+02;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 KICH 9

I: |||

Db 8 KLCH 11

RESULT 4

MLB_SCYCA
 ID MLB_SCYCA STANDARD; PRT; 18 AA.
 AC P01206;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Melanotropin beta (Beta-MSH).
 OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 CC Elasmobranchii; Galeomorphii; Galeoidea; Carchariniiformes;
 CC Scyllorhinidae; Scyllorhinus.
 OX NCBI_TaxID=7830;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=75113445; PubMed=4452470;
 RA Love R.M., Pickering B.T.;
 RT "A beta-MSH in the pituitary gland of the spotted dogfish
 RT (Scyllorhinus canicula): isolation and structure.";
 RL Gen. Comp. Endocrinol. 24:398-404(1974).
 CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
 DR PIR; A01470; MDPFEC.
 KW Hormone.
 SQ SEQUENCE 18 AA; 2195 MW; 353C93A30C314DF4 CRC64;

Query Match 24.5%; Score 24; DB 1; Length 18;

Best Local Similarity 44.4%; Pred. No. 6.2e+02;

Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 10 QIYYFGQF 18

I: |||

Db 2 ZIBYKMGHF 10

RESULT 5

ID XYCL_ACIGB STANDARD; PRT; 23 AA.
 AC P46365;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Benzaldehyde dehydrogenase [NAD+] I (EC 1.2.1.28) (Fragment).
 OS Acinetobacter genosomp. 11.
 CC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 CC Acinetobacter.
 OX NCBI_TaxID=106649;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=ATCC 11171 / NCIB 8250 / CIP 63.46 / B94;
 RX MEDLINE=91113163; PubMed=1989592;
 RA Chalmers R.M., Keen J.N., Fewson C.A.;
 RT "Comparison of benzyl alcohol dehydrogenases and benzaldehyde
 RT dehydrogenases from the benzyl alcohol and mandelate pathways in
 RT Acinetobacter calcoaceticus and from the TOL-plasmid-encoded toluene
 RT pathway in Pseudomonas putida. N-terminal amino acid sequences, amino
 RT acid compositions and immunological cross-reactions.";
 RL Biochem. J. 273:99-107(1991).
 CC -1- CATALYTIC ACTIVITY: Benzaldehyde + NAD(+) + H(2)O - benzoate +
 CC NADH.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
 DR InterPro; IPR02086; Aldehyde dehydr.
 DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; PARTIAL.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; PARTIAL.
 KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.
 FT NON_TER 23 23
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2750 MW; 50496D4930DF9E40 CRC64;

Query Match 24.5%; Score 24; DB 1; Length 23;

Best Local Similarity 55.6%; Pred. No. 7.8e+02;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 LQAKICHOI 11

I: |||

Db 3 IQTKIIEOI 11

RESULT 6

ID CXOB_CONMA STANDARD; PRT; 25 AA.
 AC P05485;
 DT 01-NOV-1998 (Rel. 09, Created)
 DT 01-NOV-1998 (Rel. 09, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Omega-conotoxin MV1B (SNX-159).
 OS Conus magus (Magus cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=6492;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=87299637; PubMed=2441741;
 RA Oliveira B.M., Cruz L.J., de Santos V., Lecheminant G.W., Griffin D.,
 RA Zeikus R.D., McIntosh J.M., Galyean R., Varga J., Gray W.R.,
 RA Rivier J.E.;
 RT "Neuronal calcium channel antagonists. Discrimination between calcium
 channel subtypes using omega-conotoxin from Conus magus venom.";
 RL Biochemistry 26:2086-2090(1987).
 CC -1- FUNCTION: Omega-conotoxins act at presynaptic membranes, they bind
 and block voltage-sensitive calcium channels (VSCC).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -1- SIMILARITY: BELONGS TO THE O-SUPERFAMILY OF CONOTOXINS. OMEGA-TYPE
 FAMILY.
 DR PIR; B34115;
 DR PIR; JH0701; JH0701.
 DR HSP; P05484; IMVI.
 KW Presynaptic neurotoxin; Neurotoxin; Toxin; Calcium channel inhibitor;
 KW Amidation.
 FT DISULFID 1 16
 FT DISULFID 8 20
 FT DISULFID 15 25
 FT MOD_RES 25 25
 SQ SEQUENCE 25 AA; 2626 MW; B4B9CE5EFAA3734D CRC64;

Query Match 24.5%; Score 24; DB 1; Length 25;
 Best Local Similarity 50.0%; Pred. No. 8.5e+02;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 8 CHQIQY 13

Db 8 CHRTSY 13

RESULT 7

ID SMBP_RAT STANDARD; PRT; 25 AA.
 AC P80968;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE SM-11044 binding protein (Fragments).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE.
 RX STRAIN=Wistar;
 RX MEDLINE=97407910; PubMed=9261134;
 RA Sugawara T., Matsuzaki-Fujita M., Guillaume J.-L., Camoin L.,
 RA Morooka S., Strosberg A.D.;
 RT "Characterization of a novel Iodocyanopindolol and SM-11044 binding
 protein, which may mediate relaxation of depolarized rat colon
 tonus.";
 RL J. Biol. Chem. 272:21244-21252(1997).
 CC -1- FUNCTION: MAY MEDIATES RELAXATION OF DEPOLARIZED COLON TONUS.

CC IT BINDS IODOCYANOPINDOLOL AND SM-11044.
 CC -1- MISCELLANEOUS: THE ORDER OF THE FRAGMENTS IS UNKNOWN.

FT NON_TER 1 1 OR Y.
 FT UNSURE 6 6
 FT NON_CONS 18 19
 FT NON_TER 25 25

SQ SEQUENCE 25 AA; 3177 MW; D14F0CB9B778C2CB CRC64;

Query Match 24.5%; Score 24; DB 1; Length 25;
 Best Local Similarity 57.1%; Pred. No. 8.5e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 12 QYFGQF 18

Db 15 QFYFPX 21

RESULT 8

MLB_SQUAC STANDARD; PRT; 16 AA.
 AC P01207;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Melanotropin beta.
 OS Squalus acanthias (Spiny dogfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
 OX NCBI_TaxID=7797;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=75127390; PubMed=4375978;
 RA Bennett H.P.J., Lowry P.J., McMartin C., Scott A.P.;
 RT "Structural studies of alpha-melanocyte-stimulating hormone and a
 novel beta-melanocyte-stimulating hormone from the neurointermediate
 lobe of the pituitary of the dogfish Squalus acanthias.";
 RL Biochem. J. 141:439-444(1974).
 CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
 DR PIR; A01471; MTDFBS.
 KW Hormone.
 SQ SEQUENCE 16 AA; 1939 MW; 993AF43C8A640A0E CRC64;

Query Match 23.5%; Score 23; DB 1; Length 16;
 Best Local Similarity 66.7%; Pred. No. 8.2e+02;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 13 YVFGQF 18

Db 5 YFQGHF 10

RESULT 9

TL18_SPIOL STANDARD; PRT; 20 AA.
 AC P82536;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Thylakoid luminal 18 kDa protein (p18) (Fragment).
 OS Spinacia oleracea (Spinach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Caryophyllaceae; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE.
 RA Kieselbach T., Bystedt M., Schroeder W.P.;
 RL Submitted (MAY-2000) to the SWISS-PROT data bank.
 CC -1- SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen.
 KW Chloroplast; Thylakoid.
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2192 MW; 752C21963F49FA64 CRC64;

Query Match 23.5%; Score 23; DB 1; Length 20;
 Best Local Similarity 27.3%; Pred. No. 1e+03;
 Matches 3; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 3 LOAKICHOIQY 13
 |||:
 Db 7 LOSKVTNKVWF 17

RESULT 10
 ANDT_ANDAU
 ID ANDT_ANDAU STANDARD; PRT; 25 AA.
 AC P56684; P81616;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Androctonus.
 OS Androctonus australis hector (Sahara scorpion).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Butthoidea; Butthidae; Androctonus.
 OX NCBI_TaxID=70175;
 RN [1]
 RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
 RC TISSUE=Hemolymph;
 RX MEDLINE=97094646; PubMed=8939880;
 RA Ehret-Sabatier L., Loew D., Goffion M., Fehlbaum P., Hoffmann J.A.,
 van Dorsseleer A., Bulet P.;
 RT "Characterization of novel cysteine-rich antimicrobial peptides from
 scorpion blood.";
 RL J. Biol. Chem. 271:29537-29544(1996).
 RN [2]

RP SYNTHESIS OF D-AMINO ACID ENANTOMER, AND CHARACTERIZATION.
 RX MEDLINE=20115101; PubMed=10642525;
 RA Hetru C., Letellier L., Oren Z., Hoffmann J.A., Shai Y.;
 RT "Androctonin, a hydrophilic disulphide-bridged non-haemolytic
 anti-microbial peptide; a plausible mode of action.";
 RL Biochem. J. 345:653-664(2000).
 RN [3]
 RP STRUCTURE BY NMR.
 RX MEDLINE=20025109; PubMed=10563585;
 RA Mandard N., Sy D., Maufrais C., Bonmatin J.M., Bulet P., Hetru C.,
 Vovelle F.;
 RT "Androctonin, a novel antimicrobial peptide from scorpion Androctonus
 australis: solution structure and molecular dynamics simulations in
 the presence of a lipid monolayer.";
 RL J. Biomol. Struct. Dyn. 17:367-380(1999).
 CC -1- FUNCTION: ACTIVE AGAINST BOTH BACTERIA (GRAM-POSITIVE AND GRAM-
 NEGATIVE) AND FILAMENTOUS FUNGI. ACTS ON THE MEMBRANE OF THE
 BACTERIAL CELLS. IT DESTABILIZE A MEMBRANE BY MODIFYING ITS
 PROPERTIES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MASS SPECTROMETRY: MW=3076.7; METHOD=Electrospray.
 DR PDB; 1CZ6; 12-JAN-00.
 KW Antibiotic; Fungicide; 3D-structure.
 FT DISULFID 4 20
 FT DISULFID 10 16
 SQ SEQUENCE 25 AA; 3081 MW; D4183D6EA5AA50AA CRC64;

Query Match 23.5%; Score 23; DB 1; Length 25;
 Best Local Similarity 57.1%; Pred. No. 1.2e+03;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 QAKICHQ 10
 | | | |
 Db 6 QIKICRR 12

RESULT 11
 AL17_CARMA
 ID AL17_CARMA STANDARD; PRT; 8 AA.
 AC P81820;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 17.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
 OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
 OC Brachyura; Eubrachyura; Portunoidea; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 8 8
 FT AMIDATION (POTENTIAL).
 SQ SEQUENCE 8 AA; 858 MW; C82879D5AB46D865 CRC64;

Query Match 22.4%; Score 22; DB 1; Length 8;
 Best Local Similarity 80.0%; Pred. No. 1.1e+05;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 QYFYG 16
 |||
 Db 3 QYSFG 7

RESULT 12
 AL11_CARMA
 ID AL11_CARMA STANDARD; PRT; 9 AA.
 AC P81814;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 11.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
 OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
 OC Brachyura; Eubrachyura; Portunoidea; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 9 9
 FT AMIDATION.
 SQ SEQUENCE 9 AA; 927 MW; 832D79DCB46D861 CRC64;

Query Match 22.4%; Score 22; DB 1; Length 9;
 Best Local Similarity 80.0%; Pred. No. 1.1e+05;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 QYFYG 16
 |||
 Db 4 QYAFG 8

RESULT 13
 LECB_PSOSC
 ID LECB_PSOSC STANDARD; PRT; 14 AA.
 AC P22584;
 DT 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Basic lectin BI (Fragment).
OS Psophocarpus scandens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eursoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
OC Psophocarpus.
OX NCBI_TaxID=3890;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RA Kortt A.A.;
RT "Isolation and characterization of the lectins from the seeds of
RL Phytochemistry 27:2847-2855(1988).
CC -1- SUBUNIT: DIMER OF TWO IDENTICAL OR NEARLY IDENTICAL SUBUNITS OF
CC -1- ABOUT 32000 APPARENT MW.
CC -1- PTM: CONTAINS ABOUT 9% CARBOHYDRATE. HETEROGENEITY IN CARBOHYDRATE
CC COMPOSITION MAY RESULT IN DIFFERENT ISOLECTIN FORMS.
CC -1- SIMILARITY: WITH P.TETRAGONOLOBUS BASIC LECTINS IN N-TERMINAL
CC SEQUENCE, AA COMPOSITION, ERYTHROCYTE AND SUGAR SPECIFICITIES.
DR PIR: PA0007; PA0007.
KW Lectin; Glycoprotein.
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1732 MW; D804CE43B487C549 CRC64;

Query Match 22.4%; Score 22; DB 1; Length 14;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 11 IQYFGQF 18
| : | | |
DB 3 ISFNQNF 10

RESULT 14
TPIS.PINPS
ID TPIS.PINPS STANDARD; PRT; 17 AA.
AC P81666;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Triosephosphate isomerase, cytosolic (EC 5.3.1.1) (TIM) (Fragments).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RP SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Prigiero J.-M., Plomont C.;
RT "Separation and characterization of needle and xylem maritime pine
RT proteins";
RL Electrophoresis 20:1098-1108(1999).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate - glycerone
CC phosphate.
CC -1- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- INDUCTION: BY WATER STRESS.
CC -1- MISCELLANEOUS: IN PLANTS, THERE ARE TWO TYPES OF TPIS, CYTOSOLIC
CC AND PLASTID.
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN
CC (SPOT N139) IS: 5.9, ITS MW IS: 24 kDa.
CC -1- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
DR InterPro: IPR000652; Triophos_1smrse.
DR PROSITE: PS00171; TIM; PARTIAL.
KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
KW Pentose shunt.
FT NON_TER 1 1

FT NON_CONS 9 10
FT NON_TER 17 17
SQ SEQUENCE 17 AA; 2053 MW; 183DB41757AF13CB CRC64;

Query Match 22.4%; Score 22; DB 1; Length 17;
Best Local Similarity 20.0%; Pred. No. 1.3e+03;
Matches 2; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 7 ICHQIQYFG 16
| : | : | : |
DB 4 VCYEQLFFVG 13

RESULT 15
BTX_ATRBI
ID BTX_ATRBI STANDARD; PRT; 21 AA.
AC P80163;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bibrotoxin (BTX).
OS Atractaspis bibroni (Southern bibron's) (Stiletto snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Atractaspididae; Atractaspis.
OX NCBI_TaxID=8601;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=93106214; PubMed=8416802;
RA Becker A., Dowdle E.B., Hechler U., Kauser K., Donner P.,
RA Schleuning W.-D.;
RT "Bibrotoxin, a novel member of the endothelin/sarafotoxin peptide
RT family, from the venom of the burrowing asp Atractaspis bibroni";
RL FEBS Lett. 315:100-103(1993).
CC -1- FUNCTION: VASOCONSTRICTOR ACTIVITY. THESE TOXINS CAUSE CARDIAC
CC ARREST PROBABLY AS A RESULT OF CORONARY VASOSPASM.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE ENDOTHELIN/SARAFOTOXIN FAMILY.
DR PIR: S27039; S27039.
DR HSSP: P13208; ISRB.
DR InterPro: IPR001928; Endothlin_tox.
DR InterPro: IPR003642; Sara/bib_toxin.
DR Pfam: PF003322; endothelin; 1.
DR PRINTS: PR00365; ENDOTHELIN.
DR ProDom: PD004740; Sara/bib_toxin; 1.
DR SMART: SM00272; END; 1.
DR PROSITE: PS00270; ENDOTHELIN; 1.
KW Vasoconstrictor; Toxin.
FT DISULFID 1 15 BY SIMILARITY.
FT DISULFID 3 11 BY SIMILARITY.
SQ SEQUENCE 21 AA; 2511 MW; 83A5DFB81D036AE2 CRC64;

Query Match 22.4%; Score 22; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQ 10
| | | |
DB 15 CHQ 17

Search completed: April 23, 2003, 13:43:49
Job time : 5.75169 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:36:43 ; Search time 9.40449 Seconds
(without alignments)
183.999 Million cell updates/sec

Title: US-09-836-073-3
Perfect score: 98
Sequence: 1 AALQAKICHQIYVFCQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 4984

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	32.7	22	2 PH1359	Ig heavy chain DJ
2	30	30.6	15	2 PH0789	T-cell receptor al
3	29	29.6	25	2 A60286	heat-stable serine
4	28	28.6	23	2 PH1725	Ig heavy chain V r
5	27	27.6	12	2 S57570	T cell receptor V-
6	27	27.6	14	2 S57569	T cell receptor V-
7	27	27.6	14	2 S57638	T cell receptor V-
8	27	27.6	22	2 PH1325	Ig heavy chain DJ
9	27	27.6	23	2 PH1681	Ig heavy chain V r
10	27	27.6	23	2 PH1724	Ig heavy chain V r
11	27	27.6	24	2 S42780	relaxin - oranguta
12	27	27.6	24	2 PH1696	Ig heavy chain V r
13	27	27.6	25	2 C57001	endo-1,4-beta-xyla
14	26	26.5	12	2 PH1324	Ig heavy chain DJ
15	26	26.5	18	2 PH1368	Ig heavy chain DJ
16	26	26.5	22	2 I77373	gene N-ras protein
17	26	26.5	23	2 PH1682	Ig heavy chain V r
18	26	26.5	24	2 PT0258	Ig heavy chain CDR
19	25	25.5	12	2 G64003	hypothetical prote
20	25	25.5	21	2 S07968	T-cell receptor be
21	25	25.5	21	2 PH1730	Ig heavy chain V r
22	25	25.5	22	2 A40741	T-cell receptor be
23	25	25.5	22	2 PH1678	Ig heavy chain V r
24	25	25.5	22	2 PH1679	Ig heavy chain V r
25	25	25.5	23	2 PH1694	Ig heavy chain V r
26	25	25.5	23	2 PH1707	Ig heavy chain V r
27	25	25.5	23	2 PH1722	Ig heavy chain V r
28	25	25.5	23	2 PH1727	Ig heavy chain V r
29	25	25.5	23	2 PH1723	Ig heavy chain V r

30	25	25.5	24	2 PH1683	Ig heavy chain V r
31	25	25.5	24	2 PH1685	Ig heavy chain V r
32	25	25.5	24	2 PH1710	Ig heavy chain V r
33	25	25.5	24	2 B53524	ubiquinol-cytochro
34	25	25.5	25	2 PH1686	Ig heavy chain V r
35	25	25.5	25	2 PH1700	Ig heavy chain V r
36	24	24.5	12	2 PH0771	T-cell receptor be
37	24	24.5	18	1 MTDFBC	melanotropin beta
38	24	24.5	18	2 PH1629	Ig H chain V-D-J r
39	24	24.5	23	2 S13298	benzaldehyde dehyd
40	24	24.5	24	2 PH1698	Ig heavy chain V r
41	24	24.5	24	2 A45336	cystic fibrosis tr
42	24	24.5	25	2 JH0701	omega-conotoxin MV
43	24	24.5	25	2 PH1716	Ig heavy chain V r
44	24	24.5	25	2 A49038	Ig lambda chain V
45	23	23.5	9	2 PT0285	Ig heavy chain CRD

ALIGNMENTS

RESULT 1

PH1359
Ig heavy chain DJ region (clone C178-121) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PH1359
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor J
A:Reference number: PH1302; MUID:93094761; PMID:1460419
A:Accession: PH1359
A:Molecule type: DNA
A:Residues: 1-22 <WAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 32.7%; Score 32; DB 2; Length 22;
Best Local Similarity 44.4%; Pred. No. 70;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 8 CHQIQYVFC 16
|::|::|
Db 6 CYENYYFG 14

RESULT 2

PH0789
T-cell receptor alpha chain (E22 V-alpha-4.delta-7R) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PH0789
R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A:Title: T cell receptor genes in a series of class I major histocompatibility comp.
allelic exclusion and antigen-specific repertoire.
A:Reference number: PH0746; MUID:92078846; PMID:1836010
A:Accession: PH0789
A:Molecule type: mRNA
A:Residues: 1-15 <CAS>
A:Cross-references: EMBL:X60894
A:Experimental source: T lymphocyte
C:Keywords: T-cell receptor

Query Match 30.6%; Score 30; DB 2; Length 15;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 9 HQIQYVFC 16
|::|::|
Db 8 HGLQYVFC 15

RESULT 3
A60286
heat-stable serine proteinase (EC 3.4.21.-) - Thermomonospora fusca (strain YX) (fragment)
N:Alternate names: YX-proteinase
C:Species: Thermomonospora fusca
C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Dec-1996
C:Accession: A60286
R:Kristjansson, M.M.; Kinsella, J.E.
Int. J. Pept. Protein Res. 36, 201-207, 1990
A:Title: Heat stable proteinase from Thermomonospora fusca. Characterization as a serine
A:Reference number: A60286; MUID:91107200; PMID:2132918
A:Accession: A60286
A:Molecule type: protein
A:Residues: 1-25 <KRI>
C:Keywords: extracellular protein; hydrolase; serine proteinase

Query Match 29.6%; Score 29; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 YFQGQF 18
|||||
DB 10 YFQGNV 15

RESULT 4
PH1725
Ig heavy chain V region (clone GCC-5) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1725
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385
A:Accession: PH1725
A:Molecule type: mRNA
A:Residues: 1-23 <MCH>
A:Experimental source: B cell
A:Note: The authors translated the codon ACA for residue 13 as Ala
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin

Query Match 28.6%; Score 28; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 3.6e+02;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIQYFQGF 18
:::|||||
DB 12 ETRYVGSY 20

RESULT 5
S57570
T cell receptor V-J junctional alpha chain region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57570
R:Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argæet, V.P.
submitted to the EMBL Data Library, June 1995
A:Description: T cell receptor repertoire for a viral epitope in humans is diversified
A:Reference number: S57494
A:Accession: S57570
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-12 <BUR>
A:Cross-references: EMBL:Z49954; NID:g887488; PIDN:CAA90225.1; PID:g887489
C:Keywords: T-cell receptor

Query Match 27.6%; Score 27; DB 2; Length 12;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 QYVFG 16
|||||
DB 8 QYVFG 12

RESULT 6
S57569
T cell receptor V-J junctional alpha chain region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57569
R:Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argæet, V.P.
submitted to the EMBL Data Library, June 1995
A:Description: T cell receptor repertoire for a viral epitope in humans is diversified
A:Reference number: S57494
A:Accession: S57569
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-14 <BUR>
A:Cross-references: EMBL:Z49955; NID:g887482; PIDN:CAA90226.1; PID:g887483
C:Keywords: T-cell receptor

Query Match 27.6%; Score 27; DB 2; Length 14;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 QYVFG 16
|||||
DB 10 QYVFG 14

RESULT 7
S57638
T cell receptor V-J junctional alpha chain region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57638
R:Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argæet, V.P.
submitted to the EMBL Data Library, June 1995
A:Description: T cell receptor repertoire for a viral epitope in humans is diversified
A:Reference number: S57494
A:Accession: S57638
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-14 <BUR>
A:Cross-references: EMBL:Z49964; NID:g886676; PIDN:CAA90238.1; PID:g886677
C:Keywords: T-cell receptor

Query Match 27.6%; Score 27; DB 2; Length 14;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 QYVFG 16
|||||
DB 10 QYVFG 14

RESULT 8
PH1325
Ig heavy chain DJ region (clone C199-121), - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PH1325
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor
A:Reference number: PH1302; MUID:93094761; PMID:1460419
A:Accession: PH1325
A:Molecule type: DNA
A:Residues: 1-22 <WAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin

Query Match 27.6%; Score 27; DB 2; Length 22;
Best Local Similarity 66.7%; Pred. No. 5.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 IQYYFG 16
| | | |
Db 9 IHYVYG 14

RESULT 9
PH1681
Ig heavy chain V region (clone NP-6-12) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1681
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385
A:Accession: PH1681
A:Molecule type: mRNA
A:Residues: 1-23 <MCH>
A:Experimental source: B cell
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 27.6%; Score 27; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 5.4e+02;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIQYFGQF 18
: | | | |
Db 12 EARYVGY 20

RESULT 10
PH1724
Ig heavy chain V region (clone GCC-4) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1724
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385
A:Accession: PH1724
A:Molecule type: mRNA
A:Residues: 1-23 <MCH>
A:Experimental source: B cell
A:Note: The authors translated the codon ACA for residue 13 as Ala
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 27.6%; Score 27; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 5.4e+02;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIQYFGQF 18
: | | | |
Db 12 ETRYVGY 20

RESULT 11
S42780
relaxin - orangutan (fragment)
C:Species: Pongo pygmaeus (orangutan)
C:Date: 08-May-1995 #sequence_revision 26-Jul-1996 #text_change 16-Jul-1999
C:Accession: S42780
R:Evans, B.B.
submitted to the EMBL Data Library, November 1993
A:Reference number: S42776
A:Accession: S42780
A:Molecule type: DNA

A:Residues: 1-24 <EVA>
A:Cross-references: EMBL:Z27229; NTD:9415990; PIDN:CAA81743.1; PID:9415991
C:Genetics:
A:Gene: rlx1
C:Superfamily: insulin
C:Keywords: disulfide bond; hormone
F:1-24/Domain: relaxin chain 1B (fragment) #status predicted <RXB1>

Query Match 27.6%; Score 27; DB 2; Length 24;
Best Local Similarity 55.6%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AALQAKICH 9
: | | | |
Db 4 SALANKCCH 12

RESULT 12
PH1696
Ig heavy chain V region (clone NP-7-9) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1696
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385
A:Accession: PH1696
A:Molecule type: mRNA
A:Residues: 1-24 <MCH>
A:Experimental source: B cell
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 27.6%; Score 27; DB 2; Length 24;
Best Local Similarity 42.9%; Pred. No. 5.6e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 10 QIQYYFG 16
: | | | |
Db 12 EVAYVYG 18

RESULT 13
CS7001
endo-1,4-beta-xylanase (EC 3.2.1.8) 4 - Streptomyces roseiscleroticus (fragment)
C:Species: Streptomyces roseiscleroticus
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 23-May-1997
C:Accession: CS7001
R:Grabski, A.C.; Forrester, I.T.; Patel, R.; Jeffries, T.W.
Protein Expr. Purif. 4, 120-129, 1993
A:Title: Characterization and N-terminal amino acid sequences of beta-(1-4)endoxyla
A:Reference number: A57001; MUID:93229899; PMID:8471845
A:Accession: CS7001
A>Status: Preliminary
A:Molecule type: protein
A:Residues: 1-25 <GRA>
A:Experimental source: strain NRRL B-11019
A:Note: sequence extracted from NCBI backbone (NCBIP:130009)
C:Function:

A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
A:Pathway: xylan degradation
C:Superfamily: Streptomyces endo-1,4-beta-xylanase A; Streptomyces endo-1,4-beta-xy
C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradati

Query Match 27.6%; Score 27; DB 2; Length 25;
Best Local Similarity 71.4%; Pred. No. 5.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIQYYFG 16
| | | | |
Db 11 QSGYYFG 17

```
RESULT 14
PHI324
Ig heavy chain DJ region (clone C510-100) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PHI324
R;Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A;Title: Predominance of fetal type DJH joining in young children with B precursor lymph
A;Reference number: PHI302; MUID:93094761; PMID:1460419
A;Accession: PHI324
A;Molecule type: DNA
A;Residues: 1-12 <WAS>
C;Keywords: heterotetramer; immunoglobulin

Query Match      26.5%; Score 26; DB 2; Length 12;
Best Local Similarity 80.0%; Pred. NO. 4.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 13 YFQG 17
   |||
Db 6 YWQG 10

RESULT 15
PHI368
Ig heavy chain DJ region (clone C111-112) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PHI368
R;Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A;Title: Predominance of fetal type DJH joining in young children with B precursor lymph
A;Reference number: PHI302; MUID:93094761; PMID:1460419
A;Accession: PHI368
A;Molecule type: DNA
A;Residues: 1-18 <WAS>
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match      26.5%; Score 26; DB 2; Length 18;
Best Local Similarity 50.0%; Pred. NO. 6.3e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 11 IQYFG 16
   :||:
Db 5 MEYFG 10
```

Search completed: April 23, 2003, 13:48:53
Job time : 9.40449 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:43:09 ; Search time 9.91011 Seconds
(without alignments)
145.542 Million cell updates/sec

Title: US-09-836-073-3
Perfect score: 98
Sequence: 1 AALQAKICHQIYFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues
Total number of hits satisfying chosen parameters: 61792

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pdb:
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pdb:
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pdb:
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pdb:
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pdb:
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pdb:
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pdb:
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pdb:
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pdb:
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pdb:
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pdb:
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pdb:
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pdb:
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pdb:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98	100.0	18	9	US-09-836-073-3
2	87	88.8	18	9	US-09-836-073-1
3	87	88.8	18	9	US-09-836-073-14
4	83	84.7	17	9	US-09-836-073-13
5	81	82.7	18	9	US-09-836-073-9
6	79	80.6	18	9	US-09-836-073-11
7	79	80.6	18	9	US-09-836-073-12
8	78	79.6	18	9	US-09-836-073-10
9	76.5	78.1	19	9	US-09-836-073-16
10	75	76.5	18	9	US-09-836-073-2
11	75	76.5	18	9	US-09-836-073-4
12	72	73.5	18	9	US-09-836-073-7
13	71	72.4	18	9	US-09-836-073-8
14	65	66.3	18	9	US-09-836-073-15
15	63	64.3	18	9	US-09-836-073-5
16	56	57.1	18	9	US-09-836-073-6
17	45	45.9	16	9	US-09-836-073-19
18	39	39.8	18	9	US-09-836-073-17
19	37.5	38.3	23	9	US-10-011-585A-144

20	30.5	31.1	18	9	US-09-836-073-18	Sequence 18, Appl
21	30	30.6	19	9	US-10-153-159-51	Sequence 51, Appl
22	30	30.6	19	9	US-10-153-159-53	Sequence 53, Appl
23	30	30.6	19	9	US-10-153-159-54	Sequence 54, Appl
24	30	30.6	19	9	US-10-153-176-51	Sequence 51, Appl
25	30	30.6	19	9	US-10-153-176-53	Sequence 53, Appl
26	30	30.6	19	9	US-10-153-176-54	Sequence 54, Appl
27	29	29.6	14	8	US-08-424-5508-455	Sequence 455, App
28	29	29.6	19	9	US-10-153-159-6	Sequence 6, Appl
29	29	29.6	19	9	US-10-153-176-6	Sequence 6, Appl
30	29	29.6	20	9	US-09-986-480-269	Sequence 269, App
31	28.5	29.1	23	9	US-09-795-515-9	Sequence 9, Appl
32	28	28.6	19	9	US-10-153-159-49	Sequence 49, Appl
33	28	28.6	19	9	US-10-153-159-59	Sequence 59, Appl
34	28	28.6	19	9	US-10-153-176-49	Sequence 49, Appl
35	28	28.6	19	9	US-10-153-176-59	Sequence 59, Appl
36	28	28.6	19	9	US-09-880-748-2987	Sequence 2987, Ap
37	28	28.6	21	10	US-09-853-830-52	Sequence 52, Appl
38	27	27.6	14	10	US-09-992-800-5	Sequence 5, Appl
39	27	27.6	14	10	US-09-992-894-5	Sequence 5, Appl
40	27	27.6	18	9	US-10-084-813-139	Sequence 139, App
41	27	27.6	18	9	US-10-084-813-140	Sequence 140, App
42	27	27.6	18	9	US-10-084-813-141	Sequence 141, App
43	27	27.6	25	9	US-10-097-065-405	Sequence 405, App
44	26	26.5	9	10	US-09-753-831-25	Sequence 25, Appl
45	26	26.5	13	9	US-09-880-748-3091	Sequence 3091, Ap

ALIGNMENTS

RESULT 1
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836.073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-836-073-3

Query Match 100.0%; Score 98; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AALQAKICHQIYFGQF 18
| | | | | | | | | | | | | | | | | |
Db 1 AALQAKICHQIYFGQF 18

RESULT 2
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836.073

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match 88.8%; Score 87; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 6.7e-08;
Matches 15; Conservative 2; Mismatches 1; Indels 0;

QY 1 AALQAKICHQIQYFGQF 18
DB 1 AALEAKICHQIEYFGDF 18

RESULT 3
US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14

Query Match 88.8%; Score 87; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 6.7e-08;
Matches 15; Conservative 2; Mismatches 1; Indels 0;

QY 1 AALQAKICHQIQYFGQF 18
DB 1 AALEAKICHQIEYFGDF 18

RESULT 4
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match 84.7%; Score 83; DB 9; Length 17;
Best Local Similarity 82.4%; Pred. No. 2.9e-07;
Matches 14; Conservative 2; Mismatches 1; Indels 0;

QY 2 AALQAKICHQIQYFGQF 18
DB 1 AALEAKICHQIEYFGDF 17

RESULT 5
US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match 82.7%; Score 81; DB 9; Length 18;
Best Local Similarity 87.5%; Pred. No. 6.4e-07;
Matches 14; Conservative 2; Mismatches 0; Indels 0;

QY 1 AALQAKICHQIQYFG 16
DB 1 AALEAKICHQIEYFG 16

RESULT 6
US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match 80.6%; Score 79; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 1.4e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0;

QY 1 AALQAKICHQIQYFGQF 18
DB 1 AALEAKICHQIEYFGDF 18

RESULT 7
US-09-836-073-12

```
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match      80.6%; Score 79; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 1.4e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYYFGQF 18
Db 1 AALEAKICHQIEYFGDF 18

RESULT 8
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match      79.6%; Score 78; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 2e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYYFGQF 18
Db 1 AALEAKICHQIEYFGDF 18

RESULT 9
US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match      78.1%; Score 76.5; DB 9; Length 19;
Best Local Similarity 78.9%; Pred. No. 3.7e-06;
Matches 15; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALQAKICHQIQYYFGQF 19
Db 1 AALEAKICHQIEYFGDF 19

RESULT 10
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match      76.5%; Score 75; DB 9; Length 18;
Best Local Similarity 72.2%; Pred. No. 6.1e-06;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYYFGQF 18
Db 1 AALEAKICHQIEYFGDF 18

RESULT 11
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match      76.5%; Score 75; DB 9; Length 18;
Best Local Similarity 80.0%; Pred. No. 6.1e-06;
```


Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 QAKICHQIQYFGQF 18
:|||||:|||||
Db 4 EAKICHQIEYFGDF 18

RESULT 12

US-09-836-073-7
; Sequence 7, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 22002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-7

Query Match

Best Local Similarity 73.5%; Score 72; DB 9; Length 18;

Matches 13; Conservative 2; Mismatches 1; Indels 0;

QY 1 AALQAKICHQIQYFG 16
:|||||:|||||
Db 1 AALEAKICHQIEYQG 16

RESULT 13

US-09-836-073-8

; Sequence 8, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 22002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-8

Query Match

Best Local Similarity 72.4%; Score 71; DB 9; Length 18;

Matches 13; Conservative 2; Mismatches 3; Indels 0;

QY 1 AALQAKICHQIQYFGQF 18
:|||||:|||||
Db 1 AALEAKICHQIEQFGDF 18

RESULT 14

US-09-836-073-15

; Sequence 15, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Xenopus
US-09-836-073-15

Query Match 66.3%; Score 65; DB 9; Length 18;

Best Local Similarity 68.8%; Pred. No. 0.00026;

Matches 11; Conservative 1; Mismatches 4; Indels 0;

QY 3 LQAKICHQIQYFGQF 18
:|||||:|||||
Db 3 LDTKICEQIEYFGDF 18

RESULT 15

US-09-836-073-5

; Sequence 5, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 22002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-5

Query Match

Best Local Similarity 64.3%; Score 63; DB 9; Length 18;

Matches 10; Conservative 2; Mismatches 3; Indels 0;

QY 4 QAKICHQIQYFGQF 18
:|||||:|||||
Db 4 EQKCHQIEYFGDF 18

Search completed: April 23, 2003, 13:52:08

Job time : 9.91011 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:38:28 ; Search time 8.49438 Seconds
(without alignments)
62.349 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98
Sequence: 1 AALQAKICHQIQYFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 147762

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	87	88.8	18	4	US-09-316-630-3		Sequence 3, Appli
2	87	88.8	18	4	US-09-316-630-4		Sequence 4, Appli
3	33	33.7	14	2	US-08-637-759B-113		Sequence 113, App
4	33	33.7	14	3	US-08-871-355A-113		Sequence 113, App
5	33	33.7	14	4	US-09-201-945-113		Sequence 113, App
6	31	31.6	20	4	US-08-505-250-8		Sequence 8, Appli
7	31	31.6	20	4	US-08-505-250-8		Sequence 8, Appli
8	30	30.6	11	2	US-08-211-312-9		Sequence 9, Appli
9	30	30.6	11	3	US-08-472-285-9		Sequence 9, Appli
10	30	30.6	11	4	US-08-472-285-9		Sequence 9, Appli
11	29	29.6	14	4	US-08-469-260A-455		Sequence 455, App
12	28.5	29.1	21	1	US-08-447-411-38		Sequence 38, Appli
13	28.5	29.1	23	2	US-08-303-569B-9		Sequence 9, Appli
14	28.5	29.1	25	4	US-08-737-629-8		Sequence 8, Appli
15	28	28.6	15	1	US-08-221-581-1		Sequence 1, Appli
16	28	28.6	15	5	PCT-US95-04018-72		Sequence 72, Appli
17	28	28.6	21	2	US-08-825-349-1		Sequence 1, Appli
18	28	28.6	22	6	5281520-43		Patent No. 5281520
19	27	27.6	7	1	US-08-166-930-15		Sequence 15, Appli
20	27	27.6	7	2	US-08-727-045A-15		Sequence 15, Appli
21	27	27.6	7	4	US-09-408-172-15		Sequence 15, Appli
22	27	27.6	9	4	US-08-747-599A-12		Sequence 12, Appli
23	27	27.6	12	4	US-08-737-841-16		Sequence 16, Appli
24	27	27.6	14	2	US-08-726-464B-47		Sequence 47, Appli
25	27	27.6	17	1	US-08-554-612C-44		Sequence 44, Appli
26	27	27.6	19	6	5464756-40		Patent No. 5464756
27	27	27.6	24	1	US-08-443-568B-4		Sequence 4, Appli

28	27	27.6	24	2	US-08-483-476-2	Sequence 2, Appli
29	27	27.6	24	2	US-08-353-476-89	Sequence 89, Appli
30	27	27.6	24	2	US-08-484-219-2	Sequence 2, Appli
31	27	27.6	24	4	US-09-158-706-2	Sequence 4, Appli
32	27	27.6	24	5	PCT-US94-06997-4	Patent No. 5464756
33	27	27.6	24	6	5464756-2	Sequence 4, Appli
34	27	27.6	25	1	US-08-453-289-4	Sequence 72, Appli
35	27	27.6	25	1	US-08-614-935-72	Sequence 73, Appli
36	27	27.6	25	1	US-08-614-935-73	Sequence 72, Appli
37	27	27.6	25	3	US-08-574-086-4	Sequence 73, Appli
38	27	27.6	25	3	US-09-130-287-72	Patent No. 5424218
39	27	27.6	25	3	US-09-130-287-73	Sequence 94, Appli
40	27	27.6	25	6	5424218-1	Sequence 47, App
41	26	26.5	14	2	US-08-433-133-94	Sequence 29, Appli
42	26	26.5	14	4	US-08-469-260A-472	Sequence 254, App
43	26	26.5	15	4	US-08-743-168B-29	Sequence 29, Appli
44	26	26.5	15	4	US-09-009-953-254	Sequence 29, Appli
45	26	26.5	15	5	PCT-US96-10435-29	

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match 88.8%; Score 87; DB 4; Length 18;
Best Local Similarity 83.3%; Pred. No. 4.9e-08;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 AALQAKICHQIQYFGQF 18
Db 1 AALEAKICHQIEYFGDF 18
|||||:|||||:|||||

RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

; PRIOR FILING DATE: 1997-10-06
 ; PRIOR APPLICATION NUMBER: 08/321,427
 ; PRIOR FILING DATE: 1994-10-11
 ; PRIOR APPLICATION NUMBER: 60/086,527
 ; PRIOR FILING DATE: 1998-05-22
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.1.
 ; SEQ ID NO 4
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
 ; OTHER INFORMATION: this peptide is biotinylated
 US-09-316-630-4

Query Match 88.8%; Score 87; DB 4; Length 18;
 Best Local Similarity 83.3%; Pred. No. 4.9e-08;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AALQAKICHQIYYFGQF 18
 Db 1 AALEAKICHQIYYFGDF 18

RESULT 3
 US-08-637-759B-113
 ; Sequence 113, Application US/08637759B
 ; Patent No. 5876931
 ; GENERAL INFORMATION:
 ; APPLICANT: David William Holden
 ; TITLE OF INVENTION: Identification of Genes
 ; NUMBER OF SEQUENCES: 501
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Patrea L. Pabst
 ; STREET: 2800 One Atlantic Center
 ; STREET: 1201 West Peachtree Street
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: USA
 ; ZIP: 30309-3450
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/637,759B
 ; FILING DATE: 03-MAY-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/GB95/02875
 ; FILING DATE: 11-DEC-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pabst, Patrea L.
 ; REGISTRATION NUMBER: 31,284
 ; REFERENCE/DOCKET NUMBER: RPMS 101
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (404) 873-8794
 ; TELEFAX: (404) 873-8795
 ; INFORMATION FOR SEQ ID NO: 113:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 14 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 US-08-637-759B-113

Query Match 33.7%; Score 33; DB 2; Length 14;
 Best Local Similarity 41.7%; Pred. No. 24;

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 Qy 7 ICHQIQYYFGQF 18
 Db 1 VCNNIQYSGHY 12

RESULT 4
 US-08-871-355A-113
 ; Sequence 113, Application US/08871355A
 ; Patent No. 6015669
 ; GENERAL INFORMATION:
 ; APPLICANT: David William Holden
 ; TITLE OF INVENTION: Identification of Genes
 ; NUMBER OF SEQUENCES: 501
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Patrea L. Pabst
 ; STREET: 2800 One Atlantic Center
 ; STREET: 1201 West Peachtree Street
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: USA
 ; ZIP: 30309-3450
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/871,355A
 ; FILING DATE: 09-JUN-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/GB95/02875
 ; FILING DATE: 11-DEC-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pabst, Patrea L.
 ; REGISTRATION NUMBER: 31,284
 ; REFERENCE/DOCKET NUMBER: RPMS 101 CON
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (404) 873-8794
 ; TELEFAX: (404) 873-8795
 ; INFORMATION FOR SEQ ID NO: 113:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 14 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 US-08-871-355A-113

Query Match 33.7%; Score 33; DB 3; Length 14;
 Best Local Similarity 41.7%; Pred. No. 24;
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 7 ICHQIQYYFGQF 18
 Db 1 VCNNIQYSGHY 12

RESULT 5
 US-09-201-945-113
 ; Sequence 113, Application US/09201945
 ; Patent No. 6342215
 ; GENERAL INFORMATION:
 ; APPLICANT: David William Holden
 ; TITLE OF INVENTION: Identification of Genes
 ; NUMBER OF SEQUENCES: 501
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Patrea L. Pabst
 ; STREET: 2800 One Atlantic Center

STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/201,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/637,759
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPMS 101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-201-945-113

Query Match 33.7%; Score 33; DB 4; Length 14;
Best Local Similarity 41.7%; Pred. No. 24;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 ICHQIQYFGQF 18
Db 1 VCNNIQYSGHY 12

RESULT 6
US-08-505-250-8
; Sequence 8, Application US/08505250
; Patent No. 6183983
; GENERAL INFORMATION:
; APPLICANT: Sato, Haruya
; APPLICANT: Yamamoto, Keiji
; APPLICANT: Suzuki, Kokichi
; APPLICANT: Ikeda, Masahiro
; APPLICANT: Taniguchi, Makoto
; TITLE OF INVENTION: PROTEIN MODIFICATION METHOD
; FILE REFERENCE: 110-511
; CURRENT APPLICATION NUMBER: US/08/505,250
; CURRENT FILING DATE: 1995-11-29
; EARLIER APPLICATION NUMBER: PCT/JP95/00298
; EARLIER FILING DATE: 1995-02-27
; EARLIER APPLICATION NUMBER: JP 198187/94
; EARLIER FILING DATE: 1994-08-23
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-08-505-250-8

Query Match 31.6%; Score 31; DB 4; Length 20;
Best Local Similarity 55.6%; Pred. No. 73;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIQYY 14
Db 3 KMCPLQOQY 11

RESULT 7
US-08-505-250-8
; Sequence 8, Application US/08505250
; Patent No. 6322996
; GENERAL INFORMATION:
; APPLICANT: Sato, Haruya
; APPLICANT: Yamamoto, Keiji
; APPLICANT: Suzuki, Kokichi
; APPLICANT: Ikeda, Masahiro
; APPLICANT: Sakagami, Masahiro
; APPLICANT: Taniguchi, Makoto
; TITLE OF INVENTION: PROTEIN MODIFICATION METHOD
; FILE REFERENCE: 110-511
; CURRENT APPLICATION NUMBER: US/08/505,250
; CURRENT FILING DATE: 1995-11-29
; PRIOR APPLICATION NUMBER: PCT/JP95/00298
; PRIOR FILING DATE: 1995-02-27
; PRIOR APPLICATION NUMBER: JP 198187/94
; PRIOR FILING DATE: 1994-08-23
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-08-505-250-8

Query Match 31.6%; Score 31; DB 4; Length 20;
Best Local Similarity 55.6%; Pred. No. 73;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIQYY 14
Db 3 KMCPLQOQY 11

RESULT 8
US-08-211-312-9
; Sequence 9, Application US/08211312
; Patent No. 5986051
; GENERAL INFORMATION:
; APPLICANT: LABIGNE, AGNES
; APPLICANT: CUSSAC, VALERIE
; APPLICANT: FERRERO, RICHARD
; TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
; TITLE OF INVENTION: FOR THE REGULATION AND MATURATION OF UREASE AND THEIR USE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
US-08-505-250-8

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/211,312
;; FILING DATE: 01-JUL-1994
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: FR 91 12198
;; FILING DATE: 03-OCT-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/FR92/00921
;; FILING DATE: 02-OCT-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Oblon, No. 5986051man F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 660-075-0XPCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 413-3000
;; TELEFAX: (703) 413-2220
;; TELEX: 248855 OPAT UR
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 11 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
US-08-211-312-9

Query Match 30.6%; Score 30; DB 2; Length 11;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 AKICHOI 11
||||:|
Db 1 AKICVEI 7

RESULT 9

US-08-472-285-9
; Sequence 9, Application US/08472285
; Patent No. 6027878
; GENERAL INFORMATION:
; APPLICANT: LABIGNE, AGNES
; APPLICANT: CUSSAC, VALERIE
; APPLICANT: FERRERO, RICHARD
; TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
; TITLE OF INVENTION: FOR THE REGULATION AND MATURATION OF UREASE AND THEIR USE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,285
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/211,312
; FILING DATE: 01-JUL-1994
; APPLICATION NUMBER: FR 91 12198
; FILING DATE: 03-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR92/00921
; FILING DATE: 02-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 6027878man F.

;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 660-075-0XPCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 413-3000
;; TELEFAX: (703) 413-2220
;; TELEX: 248855 OPAT UR
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 11 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
US-08-472-285-9

Query Match 30.6%; Score 30; DB 3; Length 11;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 AKICHOI 11
||||:|
Db 1 AKICVEI 7

RESULT 10

US-08-472-929-9
; Sequence 9, Application US/08472929
; Patent No. 6271017
; GENERAL INFORMATION:
; APPLICANT: LABIGNE, AGNES
; APPLICANT: CUSSAC, VALERIE
; APPLICANT: FERRERO, RICHARD
; TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
; TITLE OF INVENTION: FOR THE REGULATION AND MATURATION OF UREASE AND THEIR USE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,929
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/211,312
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR92/00921
; FILING DATE: 02-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 6271017man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-075-0XPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-472-929-9

Query Match 30.6%; Score 30; DB 4; Length 11;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 5 AKICHIQ 11
| | | | |
Db 1 AKICYEI 7

RESULT 11
US-08-469-260A-455
; Sequence 455, Application US/08469260A
; Patent No. 6451578
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUERHOFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,260A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 455:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-469-260A-455

Query Match 29.6%; Score 29; DB 4; Length 14;
Best Local Similarity 83.3%; Pred. No. 1.le+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 AKICHIQ 10
| | | | |
Db 3 AMICHQ 8

RESULT 12
US-08-447-411-38
; Sequence 38, Application US/08447411

; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, NO. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Naja naja
; US-08-447-411-38

Query Match 29.1%; Score 28.5; DB 1; Length 21;
Best Local Similarity 43.8%; Pred. No. 2e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

OY 2 ALQAKICHQIYFQG 17
| | | | |
Db 1 ALRLKICR---YLGE 13

RESULT 13
US-08-303-569B-9
; Sequence 9, Application US/08303569B
; Patent No. 5859205
; GENERAL INFORMATION:
; APPLICANT: Adair, John R.
; APPLICANT: Athwal, Diljeet S.
; APPLICANT: Emage, John S.
; TITLE OF INVENTION: Humanised Antibodies
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5859205r1s
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/303.569B
; APPLICATION NUMBER: US/08/303.569B
; FILING DATE: 07-SEP-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yanko
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CARP-0032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-303-569B-9.

Query Match 29.1%; Score 28.5; DB 2; Length 23;
Best Local Similarity 53.8%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

Qy 8 CHQIQ---YFGQ 17
| | | | |
Db 3 CQYQSLPYTFQ 15

RESULT 14
US-08-737-629-8
; Sequence 8, Application US/08737629
; Patent No. 6190886
; GENERAL INFORMATION:
; APPLICANT: Hoppe, Hans-Jurgen
; APPLICANT: Reid, Kenneth BM
; TITLE OF INVENTION: Trimerising polypeptides, their manufacture
; TITLE OF INVENTION: and use.
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
; STREET: 8th Floor, 1100 No. 6190886th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737.629
; FILING DATE: 10-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/01104
; FILING DATE: 16-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9409768.0
; FILING DATE: 16-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ms Mary J Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-737-629-8

Query Match 29.1%; Score 28.5; DB 4; Length 25;
Best Local Similarity 41.2%; Pred. No. 2.4e+02;
Matches 7; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

Qy 2 ALQAKICHQIQYFGQF 18
| | | : : | : | : | :
Db 10 ALQGVQH-LQAFSQY 25

RESULT 15
US-08-221-581-1
; Sequence 1, Application US/08221581
; Patent No. 5506340
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; TITLE OF INVENTION: Tumor Necrosis Factor Inhibitors
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5506340ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: United States of America
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221.581
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: CCOR-0186
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; US-08-221-581-1

Query Match 28.6%; Score 28; DB 1; Length 15;
Best Local Similarity 25.0%; Pred. No. 1.7e+02;
Matches 3; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 7 ICHQIQYFGQF 18
: | : | : | : | : | :
Db 2 LAHEVQLFSSQY 13

Search completed: April 23, 2003, 13:50:25
Job time : 9.49438 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:28:18 ; Search time 27.2022 Seconds
(without alignments)
88.173 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98

Sequence: 1 AALQAKICHQIYFQGQF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 320064

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	88.8	18	21	RAY52200 Human la autoantig
2	64	65.3	21	14	AA43394 La/SSB epitope 17.
3	37.5	38.3	23	23	ABB79195 Human prostate spe
4	34	34.7	10	23	ABB84046 Transglutaminase i
5	31	31.6	21	23	ABG66356 IgE Fcεpsilon RI b
6	30	30.6	9	23	ABG92267 PHOR1-F5D6 peptide
7	30	30.6	9	23	AAU92288 PHOR1-F5D6 peptide
8	30	30.6	9	23	AAU92398 PHOR1-F5D6 peptide
9	30	30.6	9	23	AAU92773 PHOR1-F5D6 peptide
10	30	30.6	9	23	AAU92875 PHOR1-F5D6 peptide

11	30	30.6	10	23	AAU92318
12	30	30.6	10	23	AAU92332
13	30	30.6	10	23	AAU92548
14	30	30.6	10	23	AAU92639
15	30	30.6	10	23	AAU92744
16	30	30.6	10	23	AAU92821
17	30	30.6	10	23	AAU92931
18	30	30.6	10	23	AAU92942
19	30	30.6	11	14	AA34400
20	30	30.6	14	22	AAW97227
21	30	30.6	14	22	AAW973217
22	30	30.6	16	22	AAW98909
23	30	30.6	23	11	AAW07745
24	30	30.6	23	22	AAU00924
25	29	29.6	9	20	AAW10410
26	29	29.6	11	20	AAW82311
27	29	29.6	12	21	AAW95255
28	29	29.6	12	21	AAW93796
29	29	29.6	13	19	AAW54030
30	29	29.6	14	21	AAW09328
31	29	29.6	15	22	AAW73220
32	29	29.6	19	10	AAW90453
33	29	29.6	20	21	AAW56175
34	29	29.6	21	23	ABG66506
35	29	29.6	21	23	AAW88111
36	29	29.6	21	23	AAW89824
37	29	29.6	23	17	AAW91029
38	29	29.6	25	21	AAW95957
39	28	28.6	9	20	AAW42082
40	28	28.6	9	23	ABP47453
41	28	28.6	10	22	AAW85228
42	28	28.6	10	22	AAW85230
43	28	28.6	10	23	AAE22202
44	28	28.6	11	23	ABP47593
45	28	28.6	18	19	AAW62676

ALIGNMENTS

RESULT 1
AAW52200
ID AAW52200 standard; peptide: 18 AA.

AC AAW52200;

DT 14-MAR-2000 (first entry)

DE Human la autoantigen peptide (LAP).

XX La autoantigen; LAP; internal ribosome entry site; IRES; translation;
XX viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
KW coronavirus; hepatitis virus; rhadovirus; adenovirus; coxsackie virus;
KW parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
KW vesicular stomatitis virus.

OS Homo sapiens.

PN WO961613-AZ.

XX 02-DEC-1999.

XX 21-MAY-1999; 99WO-US11281.

XX 22-MAY-1998; 98US-0086527.

PA (REGC) UNIV CALIFORNIA.

PI Das S, Dasgupta A;

DR WPI; 2000-062712/05.

PT New yeast inhibitory peptide useful for inhibiting viral protein
 PT translation and replication -
 XX
 PS Claim 5; Page 57; 81pp; English.
 XX
 CC This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (RNA see AA245200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A or C
 CC viruses, rhabdovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.

XX Sequence 18 AA;
 SQ

Query Match 88.8%; Score 87; DB 21; Length 18;
 Best Local Similarity 83.3%; Pred. No. 3.1e-07;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
 |||:|||||:|||||
 DB 1 AALEAKICHQIEYFGDF 18

RESULT 2
 AAR43394
 ID AAR43394 standard; peptide; 21 AA.

XX
 AC AAR43394;

DT 12-MAY-1994 (first entry)

DE La/SSB epitope 17.

XX Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

XX Homo sapiens.

XX WO9321223-A.

XX 28-OCT-1993.

XX 13-APR-1993; 93WO-0503484.

XX 13-APR-1992; 92US-0867819.

XX (OKLA) UNIV OKLAHOMA STATE.

XX Harley JB;

XX WPI; 1993-351658/44.

XX New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus

XX Claim 1; Page 30; 43pp; English.

XX The sequences given in AAR43391-562 are linear epitopes which are

CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polypyridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.

XX Sequence 21 AA;
 SQ

Query Match 65.3%; Score 64; DB 14; Length 21;
 Best Local Similarity 83.3%; Pred. No. 0.0018;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIQYFGQF 18

DB 1 ICHQIEYFGDF 12

RESULT 3

ABB79195

ID ABB79195 standard; Protein; 23 AA.

XX AC ABB79195;

XX 08-AUG-2002 (first entry)

XX Human prostate specific protein sequence SEQ ID NO:144.

XX Human; prostate specific gene; prostate specific protein; PSG; PSP;
 KW prostate cancer.

XX Homo sapiens.

XX WO200236808-A2.

XX 10-MAY-2002.

XX 05-NOV-2001; 2001WO-US47283.

XX 03-NOV-2000; 2000US-245740P.

XX (DIAD-) DIADEXUS INC.

XX Sun Y, Recipon H, Chen S, Liu C;

XX WPI; 2002-471506/50.

XX New prostate-specific nucleic acids and polypeptides, useful for
 PT identifying, diagnosing, monitoring, staging, imaging, and treating
 PT prostate cancer and non-cancerous disease states in prostate tissue

XX Claim 11; Page 218; 254pp; English.

XX ABN87650 to ABN87789 represent human prostate-specific nucleic acids (I),
 CC and ABN79192 to ABN79295 represent human prostate-specific proteins (II)
 CC from the present invention. (I) and (II) have cytostatic activity. (I)
 CC can be used in gene therapy. The prostate-specific nucleic acids,
 CC polypeptides and compositions from the present invention can be used for
 CC identifying, diagnosing, monitoring, staging, imaging, and treating
 CC prostate cancer and non-cancerous disease states in prostate tissue; for

CC identifying prostate tissue; for monitoring, identifying and/or designing
 CC agonists and antagonists of the polypeptides; in gene therapy; in
 CC producing transgenic animals and cells; for producing engineered prostate
 CC tissue for treatment and research; and as elements in an array or
 CC computer program for pattern recognition of prostate disorders. The
 CC nucleic acids may be used as hybridisation probes to detect, characterise
 CC and quantify hybridising nucleic acids in, and isolate hybridising
 CC nucleic acids from, both genomic and transcript-derived nucleic acid
 CC samples.

XX SQ Sequence 23 AA;

Query Match 38.3%; Score 37.5; DB 23; Length 23;
 Best Local Similarity 53.3%; Pred. No. 35;
 Matches 8; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 ALQAKICHQI-QYVF 15
 ||| ||| :|||
 Db 6 ALQTPICHTGKQYF 20

RESULT 4
 ABB84046
 ID ABB84046 standard; peptide; 10 AA.

XX AC ABB84046;

XX DT 21-AUG-2002 (first entry)

XX DE Transglutaminase inhibitory peptide cr type #16.

XX KW Transglutaminase inhibitor; Factor XIII inhibitor; XIIIa inhibitor;
 KW ophthalmological; antiinflammatory; antirheumatic; antiarthritic;
 KW thrombolytic; neuroprotective; nootropic; antiseborrheic; dermatological;
 KW cytosolic; anti-HIV; antipsoriatic; cataract; inflammatory disease;
 KW arthritis; thrombosis; Alzheimer's disease; Huntington's chorea;
 KW acne; cancer; HIV infection; psoriasis.

XX OS Unidentified.

XX PN WO200236798-A2.

XX PD 10-MAY-2002.

XX PF 02-NOV-2001; 2001WO-EP12727.

XX PR 03-NOV-2000; 2000DE-1054687.

XX PA (NZYM-) N ZYME BIOTEC GMBH.

XX PI Fuchsbaue H, Pasternack R, Zotzel J;

XX DR WPI; 2002-444364/47.

XX PT New amino acid or peptide derivatives or analogs, are selective
 PT transglutaminase inhibitors useful e.g. for treating cataract,
 PT inflammatory diseases, rheumatoid arthritis, thrombosis, Alzheimer's
 PT disease and cancer

XX PS Disclosure; Page 13; 44pp; German.

XX CC This invention describes novel amino acid or peptide derivatives or
 CC analogues (I), containing a modified side-chain (e.g. containing a formyl
 CC group) which are transglutaminase inhibitors and Factor XIII/XIIa
 CC inhibitors. The products of the invention have ophthalmological,
 CC antiinflammatory, antirheumatic, antiarthritic, thrombolytic,
 CC neuroprotective, nootropic, antiseborrheic, dermatological, cytostatic,
 CC anti-HIV and antipsoriatic activity. (I) are transglutaminase inhibitors,
 CC especially inhibitors of crosslinking of proteins or peptides
 CC (specifically fibrin and/or alpha₂-plasmin inhibitor), incorporation of
 CC primary amines in proteins and peptides, hydrolysis of the
 CC gamma-carboxamide group of glutamine residues bound in proteins or
 CC peptides, blood factor XIII/XIIa and mammalian, human, tissue, liver,

CC brain, eye lens, keratinocyte, epidermal, prostate, plant, parasitic
 CC and/or bacterial transglutaminases. The products of the invention can be
 CC used for treating cataract, inflammatory diseases, rheumatoid arthritis,
 CC chronic arthritis, thrombosis, Alzheimer's disease, Huntington's chorea,
 CC acne, cancer (by induction of apoptosis), HIV infections and psoriasis.
 CC (I) Are targeted and specific transglutaminase inhibitors, which can
 CC inhibit a specific type of transglutaminase in the human or animal body
 CC without affecting other transglutaminases. ABB84001-ABB84049 represent
 CC transglutaminase inhibitors described in the method of the invention.

XX SQ Sequence 10 AA;

Query Match 34.7%; Score 34; DB 23; Length 10;
 Best Local Similarity 85.7%; Pred. No. 54;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 QAKICHQ 10
 ||| |||
 Db 3 QAPIC HQ 9

RESULT 5
 ABB66356
 ID ABB66356 standard; Peptide; 21 AA.

XX AC ABB66356;

XX DT 30-AUG-2002 (first entry)

XX DE IgE Fcepsilon RI binding peptide IGE120 #20.

XX KW IgE receptor; immunoglobulin; FcepsilonRI; antagonist; phage display;
 KW protein co-ordinate data; IgE-mediated disease; allergic rhinitis;
 KW asthma; allergic asthma; atopic dermatitis; urticaria-angioedema;
 KW parasitic infection; IgE myeloma; immune-related disorder;
 KW inflammatory disorder; diabetes mellitus; reperfusion injury; stroke;
 KW IgE-mediated gastrointestinal inflammatory disease; burn;
 KW immune rejection of graft; myocardial infarction; atherosclerosis;
 KW acute lung injury; haemorrhagic shock; septic shock;
 KW acute tubular necrosis; endometriosis; degenerative joint disease;
 KW pancreatitis.

XX OS Synthetic.

XX PN WO200226781-A2.

XX PD 04-APR-2002.

XX PF 26-SEP-2001; 2001WO-US30289.

XX PR 26-SEP-2000; 2000US-235353P.

XX PR 23-MAR-2001; 2001US-278540P.

XX PA (GETH) GENENTECH INC.

XX PI Lowman HB, Reynolds ME, Nakamura GR, Starovasnik MA;

XX DR WPI; 2002-444016/47.

XX PT A peptide useful for treating a IgE-mediated disease or disorder in a
 PT host e.g. allergic rhinitis, asthma, which competes with immunoglobulin
 PT E for binding to high affinity IgE receptor in an in vitro assay

XX PS Example 8; Fig 2; 328pp; English.

XX CC The invention relates to a peptide which competes with immunoglobulin
 CC (Ig) E 134 comprising a sequence (S1), for binding the high affinity IgE
 CC receptor (FcepsilonRI) in an in vitro assay and having a formula given in
 CC the specification. Also included are a fusion protein comprising the
 CC peptide, a pharmaceutical composition (C) comprising the peptide,
 CC designing a compound that mimics the three-dimensional surface
 CC structure of the peptide, a compound with a solvent accessible surface
 CC that mimics the solvent accessible surface defined by the side chains of

CC residues (R) Pro4, Phe6, Pro16, Cys3, Cys7, Cys15 and Cys19 of IGE134,
 CC a peptide with structural coordinates as given in the specification,
 CC selecting a peptide mimetic which binds to fcepsilonRI and blocks
 CC binding of IgE and a peptide mimetic which mimics the coordinates of
 CC IGE134 residues (R). (C) is useful for inhibiting the binding of IgE to
 CC high affinity IgE receptor (fcepsilonRI). Peptides of the formula given
 CC in the specification are useful for inhibiting the binding of an IgE to
 CC high affinity IgE receptor. The peptide is useful for selecting a
 CC molecule which blocks the interaction of IgE with high affinity IgE
 CC receptor. The peptide is also useful for inhibiting the activation of
 CC high affinity IgE receptor. The peptide is useful for treating an IgE-
 CC mediated disease or disorder in a host. (C) is useful in research,
 CC useful for inhibiting IgE-mediated or associated processes such as IgE-
 CC dependent activation and degranulation of mast cells and basophils, as
 CC well as consequent release of inflammatory mediators such as histamine.
 CC (C) is useful for treating allergic rhinitis, asthma (e.g. allergic
 CC asthma), atopic dermatitis, urticaria-angioedema, parasitic infection,
 CC IgE myeloma, immune-related disorders, inflammatory disorders, diabetes
 CC mellitus, IgE-mediated gastrointestinal inflammatory disease, immune
 CC rejection of grafts, reperfusion injury, stroke, myocardial infarction,
 CC atherosclerosis, acute lung injury, haemorrhagic shock, burn, septic
 CC shock, acute tubular necrosis, endometriosis, degenerative joint disease
 CC and pancreatitis. The present sequence is a peptide of the invention
 CC expressed from a phage display library.

XX Sequence 21 AA;

Query Match 31.6%; Score 31; DB 23; Length 21;

Best Local Similarity 40.0%; Pred. No. 3.5e+02;

Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIQYFG 16
 DB 2 VCPRLCYWFG 11

RESULT 6

AAU92267
 ID AAU92267 standard; peptide; 9 AA.

AC AAU92267;

DT 02-JUL-2002 (first entry)

XX PHORI-F5D6 peptide #5 tested for HLA binding.

DE Human; PHORI-All; PHORI-F5D6; prostate cancer; cytostatic.

KW Homo sapiens.

OS WO200214501-A2.

PN 21-FEB-2002.

PD 17-AUG-2001; 2001WO-US25862.

PF 17-AUG-2000; 2000US-226241P.

PR (AGEN-) AGENSYS INC.

PA Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;

PI Jakobovits A;

XX WPI; 2002-269193/31.

DR Monitoring PHORI-All/PHORI-F5D6 gene products for monitoring presence
 CC of cancer in subject, by determining status of PHORI-All/PHORI-F5D6
 CC gene products in tissue sample from subject and comparing it to normal
 CC sample

PS Claim 49; Page 175; 250pp; English.

XX

CC The present invention relates to the isolation of novel human genes
 CC designated PHORI-All and PHORI-F5D6 and their encoded proteins. The
 CC gene encoding PHORI-All maps to chromosome 1q23, and the gene encoding
 CC PHORI-F5D6 maps to chromosome 7q33-q35. The PHORI-All and PHORI-F5D6
 CC polynucleotide and polypeptide sequences are useful in diagnostic and
 CC therapeutic methods, and compositions for various cancers such as
 CC prostate cancer. The sequences are useful for inhibiting the growth of
 CC cancer cells that express PHORI-All or PHORI-F5D6 and for treating
 CC cancer. The PHORI-All or PHORI-F5D6 polypeptide or a fragment thereof
 CC can be used to elicit an immune response. AAU91563-AAU92962 represent
 CC PHORI-All or PHORI-F5D6 peptides tested for HLA binding.

XX Sequence 9 AA;

Query Match 30.6%; Score 30; DB 23; Length 9;

Best Local Similarity 57.1%; Pred. No. 7.8e+05;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIQY 13

DB 3 ICHPLRY 9

RESULT 7

AAU92288
 ID AAU92288 standard; peptide; 9 AA.

AC AAU92288;

DT 02-JUL-2002 (first entry)

XX PHORI-F5D6 peptide #26 tested for HLA binding.

DE Human; PHORI-All; PHORI-F5D6; prostate cancer; cytostatic.

KW Homo sapiens.

OS WO200214501-A2.

PN 21-FEB-2002.

PD 17-AUG-2001; 2001WO-US25862.

PF 17-AUG-2000; 2000US-226241P.

PR (AGEN-) AGENSYS INC.

PA Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;

PI Jakobovits A;

XX WPI; 2002-269193/31.

DR Monitoring PHORI-All/PHORI-F5D6 gene products for monitoring presence
 CC of cancer in subject, by determining status of PHORI-All/PHORI-F5D6
 CC gene products in tissue sample from subject and comparing it to normal
 CC sample

PS Claim 49; Page 175; 250pp; English.

XX The present invention relates to the isolation of novel human genes
 CC designated PHORI-All and PHORI-F5D6 and their encoded proteins. The
 CC gene encoding PHORI-All maps to chromosome 1q23, and the gene encoding
 CC PHORI-F5D6 maps to chromosome 7q33-q35. The PHORI-All and PHORI-F5D6
 CC polynucleotide and polypeptide sequences are useful in diagnostic and
 CC therapeutic methods, and compositions for various cancers such as
 CC prostate cancer. The sequences are useful for inhibiting the growth of
 CC cancer cells that express PHORI-All or PHORI-F5D6 and for treating
 CC cancer. The PHORI-All or PHORI-F5D6 polypeptide or a fragment thereof
 CC can be used to elicit an immune response. AAU91563-AAU92962 represent
 CC PHORI-All or PHORI-F5D6 peptides tested for HLA binding.

XX Sequence 9 AA;

SQ

Query Match 30.6%; Score 30; DB 23; Length 9;
 Best Local Similarity 57.1%; Pred. No. 7.8e+05;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIQY 13
 DB 1 ICHPLRY 7

RESULT 8

AAU92398
 ID AAU92398 standard; peptide; 9 AA.

AC AAU92398;

XX 02-JUL-2002 (first entry)

XX PHOR1-F5D6 peptide #136 tested for HLA binding.

XX Human; PHOR1-All; PHOR1-F5D6; prostate cancer; cytostatic.

XX Homo sapiens.

XX WO200214501-A2.

XX 21-FEB-2002.

XX 17-AUG-2001; 2001WO-US25862.

XX 17-AUG-2000; 2000US-226241P.

XX (AGEN-) AGENSYS INC.

XX Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;

XX Jakobovits A;

XX WPI; 2002-269193/31.

XX Monitoring PHOR1-All/PHOR1-F5D6 gene products for monitoring presence
 of cancer in subject, by determining status of PHOR1-All/PHOR1-F5D6
 gene products in tissue sample from subject and comparing it to normal
 sample

XX Claim 49; Page 180; 250pp; English.

XX The present invention relates to the isolation of novel human genes
 designated PHOR1-All and PHOR1-F5D6 and their encoded proteins. The
 gene encoding PHOR1-All maps to chromosome 1q23, and the gene encoding
 PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-All and PHOR1-F5D6
 polynucleotide and polypeptide sequences are useful in diagnostic and
 therapeutic methods, and compositions for various cancers such as
 prostate cancer. The sequences are useful for inhibiting the growth of
 cancer cells that express PHOR1-All or PHOR1-F5D6 and for treating
 cancer. The PHOR1-All or PHOR1-F5D6 polypeptide or a fragment thereof
 can be used to elicit an immune response. AAU91563-AAU92962 represent
 PHOR1-All or PHOR1-F5D6 peptides tested for HLA binding.

XX Sequence 9 AA;

Query Match 30.6%; Score 30; DB 23; Length 9;
 Best Local Similarity 57.1%; Pred. No. 7.8e+05;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIQY 13
 DB 2 ICHPLRY 8

RESULT 9

AAU92773

ID AAU92773 standard; peptide; 9 AA.

XX AAU92773;

XX 02-JUL-2002 (first entry)
 XX PHOR1-F5D6 peptide #511 tested for HLA binding.

XX Human; PHOR1-All; PHOR1-F5D6; prostate cancer; cytostatic.

XX Homo sapiens.

XX WO200214501-A2.

XX 21-FEB-2002.

XX 17-AUG-2001; 2001WO-US25862.

XX 17-AUG-2000; 2000US-226241P.

XX (AGEN-) AGENSYS INC.

XX Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;

XX Jakobovits A;

XX WPI; 2002-269193/31.

XX Monitoring PHOR1-All/PHOR1-F5D6 gene products for monitoring presence
 of cancer in subject, by determining status of PHOR1-All/PHOR1-F5D6
 gene products in tissue sample from subject and comparing it to normal
 sample

XX Claim 49; Page 195; 250pp; English.

XX The present invention relates to the isolation of novel human genes
 designated PHOR1-All and PHOR1-F5D6 and their encoded proteins. The
 gene encoding PHOR1-All maps to chromosome 1q23, and the gene encoding
 PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-All and PHOR1-F5D6
 polynucleotide and polypeptide sequences are useful in diagnostic and
 therapeutic methods, and compositions for various cancers such as
 prostate cancer. The sequences are useful for inhibiting the growth of
 cancer cells that express PHOR1-All or PHOR1-F5D6 and for treating
 cancer. The PHOR1-All or PHOR1-F5D6 polypeptide or a fragment thereof
 can be used to elicit an immune response. AAU91563-AAU92962 represent
 PHOR1-All or PHOR1-F5D6 peptides tested for HLA binding.

XX Sequence 9 AA;

Query Match 30.6%; Score 30; DB 23; Length 9;
 Best Local Similarity 57.1%; Pred. No. 7.8e+05;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIQY 13
 DB 2 ICHPLRY 8

RESULT 10

AAU92875

ID AAU92875 standard; peptide; 9 AA.

XX AAU92875;

XX 02-JUL-2002 (first entry)

XX PHOR1-F5D6 peptide #613 tested for HLA binding.

XX Human; PHOR1-All; PHOR1-F5D6; prostate cancer; cytostatic.

XX Homo sapiens.

XX WO200214501-A2.

XX 21-FEB-2002.

XX 17-AUG-2001; 2001WO-US25862.


```

XX SQ Sequence 10 AA;
Query Match 30.6%; Score 30; DB 23; Length 10;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICHQIQY 13
Db 1 ICHPLRY 7

RESULT 13
AAU92548
ID AAU92548 standard; peptide; 10 AA.
XX AC
XX AAU92548;
XX DT 02-JUL-2002 (first entry)
XX PHOR1-F5D6 peptide #286 tested for HLA binding.
XX Human; PHOR1-All; PHOR1-F5D6; prostate cancer; cytostatic.
XX Homo sapiens.
XX WO200214501-A2.
XX 21-FEB-2002.
XX 17-AUG-2001; 2001WO-US25862.
XX 17-AUG-2000; 2000US-226241P.
XX (AGEN-) AGENSYS INC.
XX Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
XX Jakobovits A;
XX WPI; 2002-269193/31.
XX Monitoring PHOR1-All/PHOR1-F5D6 gene products for monitoring presence
XX of cancer in subject, by determining status of PHOR1-All/PHOR1-F5D6
XX gene products in tissue sample from subject and comparing it to normal
XX sample.
XX Claim 49; Page 186; 250pp; English.
XX The present invention relates to the isolation of novel human genes
XX designated PHOR1-All and PHOR1-F5D6 and their encoded proteins. The
XX gene encoding PHOR1-All maps to chromosome 1q23, and the gene encoding
XX PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-All and PHOR1-F5D6
XX polynucleotide and polypeptide sequences are useful in diagnostic and
XX therapeutic methods, and compositions for various cancers such as
XX prostate cancer. The sequences are useful for inhibiting the growth of
XX cancer cells that express PHOR1-All or PHOR1-F5D6 and for treating
XX cancer. The PHOR1-All or PHOR1-F5D6 polypeptide or a fragment thereof
XX can be used to elicit an immune response. AAU91563-AAU92962 represent
XX PHOR1-All or PHOR1-F5D6 peptides tested for HLA binding.
XX SQ Sequence 10 AA;
Query Match 30.6%; Score 30; DB 23; Length 10;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICHQIQY 13
Db 1 ICHPLRY 7

RESULT 14
AAU92639
ID AAU92639 standard; peptide; 10 AA.
XX AC
XX AAU92639;
XX DT 02-JUL-2002 (first entry)
XX PHOR1-F5D6 peptide #377 tested for HLA binding.
XX Human; PHOR1-All; PHOR1-F5D6; prostate cancer; cytostatic.
XX Homo sapiens.
XX WO200214501-A2.
XX 21-FEB-2002.
XX 17-AUG-2001; 2001WO-US25862.
XX 17-AUG-2000; 2000US-226241P.
XX (AGEN-) AGENSYS INC.
XX Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
XX Jakobovits A;
XX WPI; 2002-269193/31.
XX Monitoring PHOR1-All/PHOR1-F5D6 gene products for monitoring presence
XX of cancer in subject, by determining status of PHOR1-All/PHOR1-F5D6
XX gene products in tissue sample from subject and comparing it to normal
XX sample.
XX Claim 49; Page 189; 250pp; English.
XX The present invention relates to the isolation of novel human genes
XX designated PHOR1-All and PHOR1-F5D6 and their encoded proteins. The
XX gene encoding PHOR1-All maps to chromosome 1q23, and the gene encoding
XX PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-All and PHOR1-F5D6
XX polynucleotide and polypeptide sequences are useful in diagnostic and
XX therapeutic methods, and compositions for various cancers such as
XX prostate cancer. The sequences are useful for inhibiting the growth of
XX cancer cells that express PHOR1-All or PHOR1-F5D6 and for treating
XX cancer. The PHOR1-All or PHOR1-F5D6 polypeptide or a fragment thereof
XX can be used to elicit an immune response. AAU91563-AAU92962 represent
XX PHOR1-All or PHOR1-F5D6 peptides tested for HLA binding.
XX SQ Sequence 10 AA;
Query Match 30.6%; Score 30; DB 23; Length 10;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICHQIQY 13
Db 4 ICHPLRY 10

RESULT 15
AAU92744
ID AAU92744 standard; peptide; 10 AA.
XX AC
XX AAU92744;
XX DT 02-JUL-2002 (first entry)
XX PHOR1-F5D6 peptide #482 tested for HLA binding.
XX Human; PHOR1-All; PHOR1-F5D6; prostate cancer; cytostatic.
XX Homo sapiens.
XX WO200214501-A2.
XX
```

PD 21-FEB-2002.
 XX
 PF 17-AUG-2001; 2001WO-US25862.
 XX
 PR 17-AUG-2000; 2000US-226241P.
 XX
 PA (AGEN-) AGENSYS INC.
 XX
 PI Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
 PI Jakobovits A;
 XX
 DR WPI; 2002-269193/31.
 XX
 PT Monitoring PHOR1-A11/PHOR1-F5D6 gene products for monitoring presence
 PT of cancer in subject, by determining status of PHOR1-A11/PHOR1-F5D6
 PT gene products in tissue sample from subject and comparing it to normal
 PT sample -
 XX
 PS Claim 49; Page 193; 250pp; English.
 XX
 CC The present invention relates to the isolation of novel human genes
 CC designated PHOR1-A11 and PHOR1-F5D6 and their encoded proteins. The
 CC gene encoding PHOR1-A11 maps to chromosome 1q23, and the gene encoding
 CC PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-A11 and PHOR1-F5D6
 CC polynucleotide and polypeptide sequences are useful in diagnostic and
 CC therapeutic methods, and compositions for various cancers such as
 CC prostate cancer. The sequences are useful for inhibiting the growth of
 CC cancer cells that express PHOR1-A11 or PHOR1-F5D6 and for treating
 CC cancer. The PHOR1-A11 or PHOR1-F5D6 polypeptide or a fragment thereof
 CC can be used to elicit an immune response. AAU91563-AAU92962 represent
 CC PHOR1-A11 or PHOR1-F5D6 peptides tested for HLA binding.
 XX
 SQ Sequence 10 AA;

 Query Match 30.6%; Score 30; DB 23; Length 10;
 Best Local Similarity 57.1%; Pred. No. 2.4e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

 QY 7 ICHQIOY 13
 DB 3 ICHPLRY 9

Search completed: April 23, 2003, 13:42:56
 Job time : 28.2022 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:21:27 ; Search time 27.9101 Seconds
(without alignments)
132.885 Million cell updates/sec

Title: US-09-836-073-2

Perfect score: 96

Sequence: 1 AALEAQICQIIEYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:**

- 1: sp_archaea:**
- 2: sp_bacteria:**
- 3: sp_fungi:**
- 4: sp_human:**
- 5: sp_invertebrate:**
- 6: sp_mammal:**
- 7: sp_mhnc:**
- 8: sp_organelle:**
- 9: sp_phage:**
- 10: sp_plant:**
- 11: sp_rodent:**
- 12: sp_virus:**
- 13: sp_vertebrate:**
- 14: sp_unclassified:**
- 15: sp_virus:**
- 16: sp_bacteriap:**
- 17: sp_archaeap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	86.5	381	11 Q9CYB9	Q9cyb9 mus musculus
2	70	72.9	206	13 Q8QH15	Q8qh15 gallus galli
3	52	54.2	390	5 Q8R8V5	Q8r8v5 drosophila
4	50	52.1	391	2 P72382	P72382 staphylococ
5	50	52.1	391	2 P95709	P95709 staphylococ
6	50	52.1	391	16 Q9X57	Q9x57 staphylococ
7	47	49.0	411	10 Q9FL36	Q9fl36 arabidopsis
8	47	49.0	422	10 Q9A38	Q9a38 arabidopsis
9	45.5	47.4	928	10 Q9LJ02	Q9lj02 oryza sativ
10	45	46.9	396	5 Q01806	Q01806 caenorhabdi
11	45	46.9	688	5 Q9VBE6	Q9vbe6 drosophila
12	45	46.9	826	10 Q940X9	Q940x9 arabidopsis
13	44	45.8	389	10 Q8S0T8	Q8s0t8 oryza sativ
14	44	45.8	392	6 Q9GKK4	Q9gkk4 canis famli
15	44	45.8	506	10 Q9M143	Q9m143 arabidopsis
16	43	44.8	111	17 Q978R4	Q978r4 thermoplasma

17	43	44.8	119	10 Q9ZP89	Q9zp89 neurotheca
18	43	44.8	166	5 Q44678	Q44678 caenorhabdi
19	43	44.8	355	10 Q8RYB4	Q8ryb4 phytophthor
20	43	44.8	545	10 Q80567	Q80567 arabidopsis
21	43	44.8	569	10 Q49048	Q49048 arabidopsis
22	43	44.8	569	10 Q80650	Q80650 arabidopsis
23	42	43.8	159	16 Q97LL1	Q97ll1 clostridium
24	42	43.8	186	11 Q8VDM3	Q8vdm3 mus musculu
25	42	43.8	214	11 Q9CS12	Q9cs12 mus musculu
26	42	43.8	239	3 Q9P8N6	Q9p8n6 cochliobolu
27	42	43.8	328	4 Q9NW12	Q9nw12 homo sapien
28	42	43.8	342	11 Q8CUQ5	Q8cuq5 mus musculu
29	42	43.8	482	16 Q8YXQ0	Q8yxq0 brucella me
30	42	43.8	669	16 Q8XRQ1	Q8xrqi raistonia s
31	42	43.8	1379	5 Q9VAW5	Q9vaw5 drosophila
32	42	43.8	1403	5 Q9NHN6	Q9nhn6 drosophila
33	41	42.7	132	16 Q9Z6P7	Q9z6p7 listeria in
34	41	42.7	261	16 Q8XU67	Q8xu67 raistonia s
35	41	42.7	335	9 Q9AJR3	Q9ajr3 bacterioph
36	41	42.7	343	2 Q9XBI7	Q9xbi7 bacillus ce
37	41	42.7	469	16 Q66887	Q66887 aquifex aeo
38	41	42.7	505	10 Q9FOY6	Q9fqy6 capsicum an
39	41	42.7	523	10 Q94K80	Q94k80 arabidopsis
40	41	42.7	837	16 Q9PR25	Q9pr25 ureaplasma
41	41	42.7	1575	5 Q8SSN3	Q8ssn3 dictyostell
42	41	42.7	2160	5 Q17709	Q17709 caenorhabdi
43	40.5	42.2	453	5 Q9W438	Q9w438 drosophila
44	40	41.7	51	16 Q9K0R3	Q9k0r3 neisseria m
45	40	41.7	91	10 Q9AUF9	Q9auf9 brassica ca

ALIGNMENTS

RESULT 1

Q9CYB9	PRELIMINARY;	PRT;	381 AA.
ID Q9CYB9			
AC Q9CYB9			
DT 01-JUN-2001 (TremblRel. 17, Created)			
DT 01-JUN-2001 (TremblRel. 17, Last sequence update)			
DT 01-JUN-2002 (TremblRel. 21, Last annotation update)			
DE Sjogren syndrome antigen B.			
GN SSB.			
OS Mus musculus (Mouse).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX NCBI_TaxID=1090;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;			
RX MEDLINE=21085660; PubMed=11217851;			
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,			
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,			
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochi H.,			
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,			
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA Suzuki H., Toyo-Oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,			
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,			
RA Hayashizaki Y.;			
RT "Functional annotation of a full-length mouse cDNA collection.";			
RL Nature 409:685-690(2001).			
DR EMBL; AK017822; BAB30957.1; -			
DR MGD; MGI:98423; Ssb.			

DR InterPro: IPR002344; Lupus.La.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; rim; 1.
 DR PRINTS: PR00302; LUPUSLA.
 SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;

Query Match 86.5%; Score 83; DB 11; Length 381;
 Best Local Similarity 88.2%; Pred. No. 8.4e-06;
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALEAQICQIEYFGDF 18
 |||||:|||||
 Db 12 ALEAKICHQIEYFGDF 28

RESULT 2
 Q8QH15 PRELIMINARY; PRT; 206 AA.
 AC Q8QH15;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DE Untranslated region binding-protein.
 GN vbp.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA L'Ecuquer T.J., Fang H.-L.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF467897; AAL76269.1;
 SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFB90E9 CRC64;

Query Match 72.9%; Score 70; DB 13; Length 206;
 Best Local Similarity 80.0%; Pred. No. 0.00069;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 LEAQICQIEYFGD 17
 |||||:|||||
 Db 13 LESKICQIEYFGN 27

RESULT 3
 Q8T8V5 PRELIMINARY; PRT; 390 AA.
 ID Q8T8V5;
 AC Q8T8V5;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DE AT22034p.
 GN LA.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celisner S.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY075257; AAL68124.1;
 SQ SEQUENCE 390 AA; 44869 MW; 797FDE26B903C909 CRC64;

Query Match 54.2%; Score 52; DB 5; Length 390;
 Best Local Similarity 64.3%; Pred. No. 1.4;

Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 EAQICQIEYFGD 17
 | | :|||||
 Db 51 ERAIIRQVEYFGD 64

RESULT 4
 P72382 PRELIMINARY; PRT; 391 AA.
 ID P72382;
 AC P72382;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Cap8p.
 GN CAP8P.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BECKER;
 RX MEDLINE=96178981; PubMed=8606192;
 RA Sau S., Lee C.Y.;
 RT "Cloning of type 8 capsule genes and analysis of gene clusters for the
 RT production of different capsular polysaccharides in Staphylococcus
 RT aureus.";
 RL J. Bacteriol. 178:2118-2126(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BECKER;
 RX MEDLINE=97197525;
 RA Sau S., Sun J., Lee C.Y.;
 RT "Molecular characterization and transcriptional analysis of type 8
 RT capsule genes in Staphylococcus aureus.";
 RL J. Bacteriol. 179:1614-1621(1997).
 DR HSSP: P27828; 1F6D.
 DR InterPro: IPR003331; Epimerase_2.
 DR Pfam: PF02350; Epimerase_2; 1.
 DR TIGRFAMS: TIGR00236; wecB; 1.
 SQ SEQUENCE 391 AA; 44247 MW; 1E8D9FAA9BC76F0D CRC64;

Query Match 52.1%; Score 50; DB 2; Length 391;
 Best Local Similarity 72.7%; Pred. No. 3.1;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 QICQIEYFG 16
 |||:|||||
 Db 366 RICEAIEYFG 376

RESULT 5
 P95709 PRELIMINARY; PRT; 391 AA.
 ID P95709;
 AC P95709;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Cap5p.
 GN CAP5P.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-REYNOLDS, AND NEWMAN;
 RX MEDLINE=97388587; PubMed=9245821;
 RA Sau S., Bhasin N., Wann E.R., Lee J.C., Foster T.J., Lee C.Y.;
 RT "The Staphylococcus aureus allelic genetic loci for serotype 5 and 8
 RT capsule expression contain the type-specific genes flanked by common

```

RT genes."
RL Microbiology 143:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=REYNOLDS, AND NEWMAN;
RX MEDLINE=98101481; PubMed=9440531;
RA Kiser K.B., Lee J.C.;
RT "Staphylococcus aureus cap50 and cap5P genes functionally complement
RT mutations affecting enterobacterial common-antigen biosynthesis in
RT Escherichia coli.";
RL J. Bacteriol. 180:403-406(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=REYNOLDS, AND NEWMAN;
RX MEDLINE=98125727; PubMed=9466251;
RA Bhasin N., Albus A., Michon F., Livolsi P.J., Park J.-S., Lee J.C.;
RT "Identification of a gene essential for O-acetylation of the
RT Staphylococcus aureus type 5 capsular polysaccharide.";
RL Mol. Microbiol. 27:9-21(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=REYNOLDS, AND NEWMAN;
RA Bagga N., Wann E.R., Foster
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U81973; AAC46099.1; -.
DR HSSP; P27828; 1F6D.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMs; TIGR00236; weCB; 1.
SQ SEQUENCE 391 AA; 44328 MW; B0105F690DB7CF1D CRC64;

Query Match 52.1%; Score 50; DB 2; Length 391;
Best Local Similarity 72.7%; Pred. No. 3.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 QICQOIEYFG 16
Db 366 RICEAIEYFG 376

RESULT 6
Q99X57 PRELIMINARY; PRT; 391 AA.
AC Q99X57;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Capsular polysaccharide synthesis enzyme Cap5P.
GN CAPP OR SAV0164 OR SA0159.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshida K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003358; BAB56326.1; -.
DR EMBL; AP003129; BAB41379.1; -.
DR HSSP; P27828; 1F6D.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.

```

```

DR TIGRFAMs; TIGR00236; weCB; 1.
KW Complete proteome.
SQ SEQUENCE 391 AA; 44372 MW; D0DF5FA715BCCECC CRC64;

Query Match 52.1%; Score 50; DB 16; Length 391;
Best Local Similarity 72.7%; Pred. No. 3.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 QICQOIEYFG 16
Db 366 RICEAIEYFG 376

RESULT 7
Q9FL36 PRELIMINARY; PRT; 411 AA.
AC Q9FL36;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similarity to RNA-binding protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98344145; PubMed=9679202;
RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT features of the regions of 1,381,565 bp covered by twenty one
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:131-145(1998).
DR EMBL; AB010698; BAB11080.1; -.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
SQ SEQUENCE 411 AA; 45655 MW; A2EF62EB58B9B099 CRC64;

Query Match 49.0%; Score 47; DB 10; Length 411;
Best Local Similarity 53.3%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LEAQICQIEYFGD 17
Db 92 LNKIIRQVEYFSD 106

RESULT 8
Q94A38 PRELIMINARY; PRT; 422 AA.
AC Q94A38;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE AT5G46250/NPL12_3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,

```

```

RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.C., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
RT "Arabidopsis cDNA clones.",
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY050403; AAK91419.1; -.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rtm: 1.
DR PROSITE: PS0102; RRM; 1.
SQ SEQUENCE 422 AA; 46842 MW; 4EC4BBBF1E069F0E CRC64;

Query Match 49.0%; Score 47; DB 10; Length 422;
Best Local Similarity 53.3%; Pred. No. 11;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LEAQICQIQIEYFGD 17
| : : : : :
Db 103 LNKILRQVEYFSD 117
| : : : : :

RESULT 9
O9LJ02 PRELIMINARY; PRT; 928 AA.
AC Q9LJ02;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Similar to KIAA0731 protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]

SEQUENCE FROM N.A.
STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone: P0499C11.",
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF001080; BAA90356.1; -.
SQ SEQUENCE 928 AA; 103745 MW; 53FC46E24A446EB4 CRC64;

Query Match 47.4%; Score 45.5; DB 10; Length 928;
Best Local Similarity 62.5%; Pred. No. 43;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 3 LEAQICQIQIEYFGD 17
| : : : : :
Db 278 LRKILRQVEYFSGD 293
| : : : : :

RESULT 10
O01806 PRELIMINARY; PRT; 396 AA.
AC O01806;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE C44E4.4 protein.
GN C44E4.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
SEQUENCE FROM N.A.
STRAIN=BRISTOL N2;
RC STRAIN=BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
```

```

RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Sammons L., Wohldmann P., Gillam B.;
RT "The sequence of C. elegans cosmid C44E4.",
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003140; AAB54169.1; -.
DR InterPro: IPR002344; Lupus_La.
DR Pfam: PF00076; rtm: 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS0102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
SQ SEQUENCE 396 AA; 43631 MW; A2D82844AFAA3C34 CRC64;

Query Match 46.9%; Score 45; DB 5; Length 396;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAQICQIQIEYFGD 17
| : : : : :
Db 13 DQKILRQVEYFGN 26
| : : : : :

RESULT 11
Q9VBE6 PRELIMINARY; PRT; 688 AA.
AC Q9VBE6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CG5455 protein (GH15286P).
GN CG5455
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadenot L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
```


DR PROSITE: PS00627: GHMP_KINASES_ATP: 1.
 KW ATP-binding; Galactose metabolism; Kinase; Transferase.
 SQ SEQUENCE 392 AA; 42120 MW; 8FD5ECF3AB4386BF CRC64;

Query Match 45.8%; Score 44; DB 6; Length 392;

Best Local Similarity 57.1%; Pred. NO. 32;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 ALEAQICQIEYFG 15

DB 164 ARAQVCOQAEHSF 177

RESULT 15

Q9M143 PRELIMINARY; PRT; 506 AA.

AC Q9M143;

DT 01-OCT-2000 (TRENBLrel. 15, Created)

DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

DE Putative RING zinc finger protein.

GN AR4G01270.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,

RA Mayer K.F.X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

DR EMBL: AL161491; CAB80936.1; .

DR InterPro: IPR001841; Znf_fing.

DR Pfam: PF00097; zf-C3HC4; 1.

DR SMART: SM00184; RING; 1.

KW Zinc-finger.

SQ SEQUENCE 506 AA; 56518 MW; 1B06751A4E299FE8 CRC64;

Query Match 45.8%; Score 44; DB 10; Length 506;

Best Local Similarity 66.7%; Pred. NO. 41;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 AQICQIEYFG 16

DB 491 SQGCLQIEHYFG 502

Search completed: April 23, 2003, 13:32:47

Job time : 29.9101 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:27:28 ; Search time 10.3146 Seconds
(without alignments)
139.834 Million cell updates/sec

Title: US-09-836-073-2

Perfect score: 96
Sequence: 1 AALEAQICQOIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications_AA.*
- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
 - 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
 - 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	18	9	US-09-836-073-2
2	87	90.6	18	9	US-09-836-073-1
3	87	90.6	18	9	US-09-836-073-14
4	87	90.6	460	9	US-10-102-806-695
5	83	86.5	17	9	US-09-836-073-13
6	81	84.4	18	9	US-09-836-073-9
7	79	82.3	18	9	US-09-836-073-11
8	79	82.3	18	9	US-09-836-073-12
9	78	81.2	18	9	US-09-836-073-10
10	76.5	79.7	19	9	US-09-836-073-16
11	75	78.1	18	9	US-09-836-073-3
12	75	78.1	18	9	US-09-836-073-4
13	74	77.1	18	9	US-09-836-073-15
14	72	75.0	18	9	US-09-836-073-7
15	71	74.0	18	9	US-09-836-073-8
16	63	65.6	18	9	US-09-836-073-5
17	61	63.5	38	9	US-09-843-676-25
18	61	63.5	38	9	US-09-766-253-25
19	61	63.5	38	9	US-09-438-486-25

20	61	63.5	38	9	US-10-053-758-25	Sequence 25, Appl
21	61	63.5	38	9	US-10-054-295-25	Sequence 25, Appl
22	61	63.5	38	9	US-10-054-611-25	Sequence 25, Appl
23	58	60.4	16	9	US-09-836-073-19	Sequence 19, Appl
24	56	58.3	18	9	US-09-836-073-6	Sequence 6, Appl
25	50.5	52.6	37	9	US-09-843-676-24	Sequence 24, Appl
26	50.5	52.6	37	9	US-09-766-253-24	Sequence 24, Appl
27	50.5	52.6	37	9	US-09-438-486-24	Sequence 24, Appl
28	50.5	52.6	37	9	US-10-053-758-24	Sequence 24, Appl
29	50.5	52.6	37	9	US-10-054-295-24	Sequence 24, Appl
30	50.5	52.6	37	9	US-10-054-611-24	Sequence 24, Appl
31	49	51.0	39	9	US-09-843-676-26	Sequence 26, Appl
32	49	51.0	39	9	US-09-766-253-26	Sequence 26, Appl
33	49	51.0	39	9	US-09-438-486-26	Sequence 26, Appl
34	49	51.0	39	9	US-10-053-758-26	Sequence 26, Appl
35	49	51.0	39	9	US-10-054-295-26	Sequence 26, Appl
36	49	51.0	39	9	US-10-054-611-26	Sequence 26, Appl
37	45	46.9	18	9	US-09-836-073-17	Sequence 17, Appl
38	44	45.8	406	9	US-10-102-806-668	Sequence 668, App
39	43.5	45.3	18	9	US-09-836-073-18	Sequence 18, Appl
40	42	43.8	211	10	US-09-925-302-786	Sequence 786, App
41	42	43.8	487	9	US-10-023-437-19	Sequence 19, Appl
42	41	42.7	229	10	US-09-893-737-318	Sequence 318, App
43	41	42.7	382	10	US-09-893-737-36	Sequence 36, Appl
44	40	41.7	262	9	US-09-813-453A-8	Sequence 8, Appl
45	40	41.7	505	10	US-09-765-873A-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836, 073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match 100.0%; Score 96; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAQICQOIEYFGDF 18
| | | | | | | | | | | | | | | | | |
Db 1 AALEAQICQOIEYFGDF 18

RESULT 2
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836, 073

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match 90.6%; Score 87; DB 9; Length 18;
Best Local Similarity 88.9%; Pred. No. 3.3e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIQEYFGDF 18
|||||:|||||
Db 1 AALEAKICHOIEYFGDF 18

RESULT 3
US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14

Query Match 90.6%; Score 87; DB 9; Length 18;
Best Local Similarity 88.9%; Pred. No. 3.3e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIQEYFGDF 18
|||||:|||||
Db 1 AALEAKICHOIEYFGDF 18

RESULT 4
US-10-102-806-695
; Sequence 695, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 695
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-102-806-695

Query Match 90.6%; Score 87; DB 9; Length 460;
Best Local Similarity 88.9%; Pred. No. 6.9e-06;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIQEYFGDF 18
|||||:|||||
Db 63 AALEAKICHOIEYFGDF 80

RESULT 5
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match 86.5%; Score 83; DB 9; Length 17;
Best Local Similarity 88.2%; Pred. No. 1.3e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AALEAQCQIQEYFGDF 18
|||||:|||||
Db 1 AALEAKICHOIEYFGDF 17

RESULT 6
US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match 84.4%; Score 81; DB 9; Length 18;
Best Local Similarity 88.2%; Pred. No. 2.8e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIQEYFGD 17
|||||:|||||
Db 1 AALEAKICHOIEYFGD 17

RESULT 7

US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match 82.3%; Score 79; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 5.7e-06;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALEAQICQOIEYFGDF 18
|||||:|||||
Db 1 AALEAKICHQIEYQGF 18

RESULT 8

US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match 82.3%; Score 79; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 5.7e-06;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALEAQICQOIEYFGDF 18
|||||:|||||
Db 1 AALEAKICHQIEYQGF 18

RESULT 9

US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match 81.2%; Score 78; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 8.1e-06;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALEAQICQOIEYFGDF 18
|||||:|||||
Db 1 AALEAKICHQIEYQGF 18

RESULT 10

US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match 79.7%; Score 76.5; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 1.5e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALEAQICQOIEYFGDF 18
|||||:|||||
Db 1 AALEAKICHQIEYFGDF 19

RESULT 11

US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match 78.1%; Score 75; DB 9; Length 18;
Best Local Similarity 72.2%; Pred. No. 2.4e-05;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALEAQCQIEYYFGDF 18
DB 1 AALQAKICHQIEYYFGDF 18

RESULT 12
US-09-836-073-4

; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match 78.1%; Score 75; DB 9; Length 18;
Best Local Similarity 86.7%; Pred. No. 2.4e-05;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAQICQIEYYFGDF 18
DB 4 EAKICHQIEYYFGDF 18

RESULT 13

; Sequence 15, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Xenopus
US-09-836-073-15

Query Match 77.1%; Score 74; DB 9; Length 18;
Best Local Similarity 75.0%; Pred. No. 3.4e-05;
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LEAQICQIEYYFGDF 18
DB 3 LDTKICQIEYYFGDF 18

RESULT 14
US-09-836-073-7

; Sequence 7, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-7

Query Match 75.0%; Score 72; DB 9; Length 18;
Best Local Similarity 82.4%; Pred. No. 6.9e-05;
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALEAQCQIEYYFGD 17
DB 1 AALEAKICHQIEYYQGD 17

RESULT 15

US-09-836-073-8
; Sequence 8, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-8

Query Match 74.0%; Score 71; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 9.8e-05;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AALEAQCQIEYYFGDF 18
DB 1 AALEAKICHQIEQQQFGDF 18

Search completed: April 23, 2003, 13:38:19
Job time : 11.3146 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:24 ; Search time 11.2247 Seconds
(without alignments)
47.183 Million cell updates/sec

Title: US-09-836-073-2

Perfect score: .96
Sequence: 1 AALEAQICQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2.6/ptodata/1/1aa/5A-COMB.pep.*
2: /cgn2.6/ptodata/1/1aa/5B-COMB.pep.*
3: /cgn2.6/ptodata/1/1aa/6A-COMB.pep.*
4: /cgn2.6/ptodata/1/1aa/6B-COMB.pep.*
5: /cgn2.6/ptodata/1/1aa/PCTUS-COMB.pep.*
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	87	90.6	18	4	US-09-316-630-3	Sequence 3, Appl
2	87	90.6	18	4	US-09-316-630-4	Sequence 4, Appl
3	61	63.5	38	3	US-08-851-843A-25	Sequence 25, Appl
4	61	63.5	38	4	US-08-974-549A-215	Sequence 215, Appl
5	61	63.5	38	4	US-08-854-050-25	Sequence 25, Appl
6	61	63.5	38	4	US-09-430-323-25	Sequence 25, Appl
7	60	62.5	38	4	US-08-974-549A-214	Sequence 214, Appl
8	50.5	52.6	37	3	US-08-851-843A-24	Sequence 24, Appl
9	50.5	52.6	37	4	US-08-854-050-24	Sequence 24, Appl
10	50.5	52.6	37	4	US-09-430-323-24	Sequence 24, Appl
11	49	51.0	39	3	US-08-851-843A-26	Sequence 26, Appl
12	49	51.0	39	4	US-08-974-549A-216	Sequence 216, Appl
13	49	51.0	39	4	US-08-854-050-26	Sequence 26, Appl
14	49	51.0	39	4	US-09-430-323-26	Sequence 26, Appl
15	44	45.8	392	1	US-08-451-777A-33	Sequence 33, Appl
16	44	45.8	392	2	US-08-451-778A-33	Sequence 33, Appl
17	44	45.8	392	2	US-08-998-208-33	Sequence 33, Appl
18	41	42.7	29	1	US-08-451-777A-3	Sequence 3, Appl
19	41	42.7	29	2	US-08-451-778A-3	Sequence 3, Appl
20	41	42.7	29	2	US-08-998-208-3	Sequence 3, Appl
21	41	42.7	29	5	PCT-US94-10825-3	Sequence 3, Appl
22	41	42.7	29	5	PCT-US95-06743-3	Sequence 3, Appl
23	40	41.7	505	4	US-09-627-216A-12	Sequence 12, Appl
24	40	41.7	505	4	US-09-126-420A-22	Sequence 22, Appl
25	40	41.7	754	2	US-08-941-262-1	Sequence 1, Appl
26	40	41.7	755	2	US-08-941-262-3	Sequence 3, Appl
27	40	41.7	1180	3	US-09-224-024-28	Sequence 28, Appl

28	40	41.7	1180	5	PCT-US94-07902-28	Sequence 28, Appl
29	39	40.6	456	1	US-08-205-719-4	Sequence 4, Appl
30	39	40.6	456	4	US-08-431-517F-6	Sequence 6, Appl
31	39	40.6	482	4	US-08-431-517F-5	Sequence 5, Appl
32	39	40.6	482	6	5245013-2	Patent No. 5245013
33	38.5	40.1	407	1	US-08-487-823B-2	Sequence 2, Appl
34	38.5	40.1	407	2	US-08-997-040-2	Sequence 2, Appl
35	38.5	40.1	407	2	US-09-203-237-2	Sequence 2, Appl
36	38.5	40.1	410	3	US-08-948-997-2	Sequence 2, Appl
37	38.5	40.1	410	4	US-09-348-817A-2	Sequence 2, Appl
38	38.5	40.1	410	4	US-09-348-817A-3	Sequence 3, Appl
39	38	39.6	38	3	US-08-851-843A-27	Sequence 27, Appl
40	38	39.6	38	4	US-08-974-549A-217	Sequence 217, Appl
41	38	39.6	38	4	US-08-854-050-27	Sequence 27, Appl
42	38	39.6	38	4	US-09-430-323-27	Sequence 27, Appl
43	38	39.6	379	4	US-09-603-185-6	Sequence 6, Appl
44	38	39.6	556	2	US-08-505-377-1	Sequence 1, Appl
45	38	39.6	556	3	US-08-798-269-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match 90.6%; Score 87; DB 4; Length 18;
Best Local Similarity 88.9%; Pred.No. 2.7e+08;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQICQIEYFGDF 18
|||||:|||||
Db 1 AALEAKICQIEYFGDF 18

RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
; OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 90.6%; Score 87; DB 4; Length 18;
Best Local Similarity 88.9%; Pred. No. 2.7e-08;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQICQOIEYYFGDF 18
|||||:|||||
Db 1 AALEAKICHOIEYYFGDF 18

RESULT 3
US-08-851-843A-25
; Sequence 25, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-851-843A-25

Query Match 63.5%; Score 61; DB 3; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0013;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICQOIEYYFGD 17
||:|||||
Db 1 ICQOIEYYFGD 11

RESULT 4
US-08-974-549A-215
; Sequence 215, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US97/17618
;; FILING DATE: 01-OCT-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US97/17885
;; FILING DATE: 01-OCT-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph Ted
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002610US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 215:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-974-549A-215

Query Match 63.5%; Score 61; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0013;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICQIIEYFGD 17
||:|||||||
Db 1 ICEQIEYFGD 11

RESULT 5
US-08-854-050-25
; Sequence 25, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002930US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
US-08-854-050-25

Query Match 63.5%; Score 61; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0013;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICQIIEYFGD 17
||:|||||||
Db 1 ICEQIEYFGD 11

RESULT 6
US-09-430-323-25
; Sequence 25, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-00293005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match 63.5%; Score 61; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0013;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICQIEYYFGD 17
||:|||||||
Db 1 ICEQIEYYFGD 11

RESULT 7

US-08-974-549A-214
Sequence 214, Application US/08974549A
Patent No. 6166178

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.

TITLE OF INVENTION: Human Telomerase Catalytic Subunit

NUMBER OF SEQUENCES: 727

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,549A

FILING DATE: 19-NOV-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/911,312

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/912,951

FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00261005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-214

Query Match 62.5%; Score 60; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.0019;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICQIEYYFGDF 18
||:|||||||
Db 1 ICHQXEYIFGDF 12

RESULT 8

US-08-851-843A-24

Sequence 24, Application US/08851843A

Patent No. 6093809

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin

APPLICANT: Andrews, William H.

TITLE OF INVENTION: NO. 6093809el Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/851,843A

FILING DATE: 06-MAY-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

CLASSIFICATION:

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002930US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 37 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
US-08-851-843A-24

Query Match 52.6%; Score 50.5; DB 3; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.07;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICQIEYYFGDF 18
||| |||||
Db 1 ICHQ-EYYFGDF 11

RESULT 9
US-08-854-050-24
; Sequence 24, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643

;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002930US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 37 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
US-08-854-050-24

Query Match 52.6%; Score 50.5; DB 4; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.07;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICQIEYYFGDF 18
||| |||||
Db 1 ICHQ-EYYFGDF 11

RESULT 10
US-09-430-323-24
; Sequence 24, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-430-323-24

Query Match 52.6%; Score 50.5; DB 4; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.07;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICQIEYIFGDF 18
II I I I I I I I I I I
Db 1 ICQI-EYIFGDF 11

RESULT 11
US-08-851-843A-26
Sequence 26, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
LENGTH: 39 amino acids

REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-0029300S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids

TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-26
Query Match 51.0%; Score 49; DB 3; Length 39;
Best Local Similarity 72.7%; Pred. No. 0.13;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 7 ICQIEYIFGDF 17
I I I I I I I I I I
Db 1 ILRQVEYIFGD 11

RESULT 12
US-08-974-549A-216
Sequence 216, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 216:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-549A-216

Query Match      51.0%; Score 49; DB 4; Length 39;
Best Local Similarity 72.7%; Pred. No. 0.13;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICQIETYYFGD 17
Db 1 ILRQVEYYFGD 11

RESULT 13
US-08-854-050-26
; Sequence 26, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-854-050-26

Query Match      51.0%; Score 49; DB 4; Length 39;
Best Local Similarity 72.7%; Pred. No. 0.13;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICQIETYYFGD 17
Db 1 ILRQVEYYFGD 11

RESULT 14
US-09-430-323-26
; Sequence 26, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
```


Search completed: April 23, 2003, 13:36:30
Job time : 12.2247 secs

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-430-323-26

Query Match 51.0%; Score 49; DB 4; Length 39;
Best Local Similarity 72.7%; Pred. No. 0.13;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 7 ICQIIEYFGD 17
| :|:|:|:|:|
DB 1 ILQVEYFGD 11

RESULT 15
US-08-451-777A-33
; Sequence 33, Application US/08451777A
; Patent No. 5789223
; GENERAL INFORMATION:
; APPLICANT: Bergsma, Derk J.
; APPLICANT: Stambolian, Dwight
; TITLE OF INVENTION: Human Galactokinase Gene
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corp./Corporate
; ADDRESSEE: Intellectual Property
; STREET: 709 Swedeland Road/UW2220
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,777A
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10825
; FILING DATE: 23-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Eagle, Alissa M.
; REGISTRATION NUMBER: 37,126
; REFERENCE/DOCKET NUMBER: P50268-1B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5364
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 392 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-451-777A-33

Query Match 45.8%; Score 44; DB 1; Length 392;
Best Local Similarity 57.1%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 2 ALEAQICQIIEYF 15
| :|:|:|:|:|
DB 164 AARQVCQAHSF 177

Result No.	Score	Query Match	Length	DB ID	Description
1	87	90.6	18	21	AAY52200
2	87	90.6	92	21	Human la autoantig
3	87	90.6	92	21	Human secreted pro
4	87	90.6	408	17	Human autoantigen
5	87	90.6	439	22	Novel human diagno
6	87	90.6	460	21	Breast and ovarian
7	87	90.6	460	23	Human ovarian anti
8	55	67.7	21	14	La/SSB epitope 17.
9	52	54.2	390	22	Drosophila melanog
10	45	46.9	688	22	Drosophila melanog
11	44	45.8	392	17	Human galactokinase

PT New yeast inhibitory peptide useful for inhibiting viral protein
 translation and replication .
 PS Claim 5; Page 57; 81pp; English.
 CC This sequence is the La autoantigen binding domain (LAP). LAP is a
 cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (RNA see AA45200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rhinovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.
 XX
 SQ Sequence 18 AA;

Query Match 90.6%; Score 87; DB 21; Length 18;
 Best Local Similarity 88.9%; Pred. No. 1.6e-07;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAICQIQIEYFGDF 18
 |||||:|||||
 Db 1 AALEAKICHQIEYFGDF 18

RESULT 2

AAAG01351
 ID AAG01351 standard; Protein; 92 AA.

XX
 AC AAG01351;

XX 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 5432.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 gene therapy; chromosome mapping.

OS Homo sapiens.

XX EF1033401-A2.

PN 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.

DR N-PSDB; AAC01357.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 13; SEQ ID 5432; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
 of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 were prepared from total human RNAs or polyA+ RNAs derived from 30

CC different tissues. EST sequences usually correspond mainly to the 3'
 untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX

SQ Sequence 92 AA;

Query Match 90.6%; Score 87; DB 21; Length 92;
 Best Local Similarity 88.9%; Pred. No. 9.6e-07;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAICQIQIEYFGDF 18
 |||||:|||||
 Db 11 AALEAKICHQIEYFGDF 28

RESULT 3

AAW03716

ID AAW03716 standard; protein; 408 AA.

XX
 AC AAW03716;

XX 12-MAR-1997 (first entry)

XX Human autoantigen La(SS-B).

XX Autoimmune disease; La autoantigen; Sjogren's syndrome;
 KW systemic lupus erythematosus; diagnosis.

OS Homo sapiens.

XX US5541291-A.

XX 30-JUL-1996.

XX 31-DEC-1984; 84US-0687908.

XX 27-MAY-1987; 87US-0054871.

XX 31-DEC-1984; 84US-0687908.

XX (UYDU-) UNIV DUKE.

XX Keene JD;

XX WPI; 1996-362015/36.

XX Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
 PT overlap syndrome - useful for diagnosis and treatment of autoimmune
 PT diseases

XX Disclosure; Columns 15-16; 21pp; English.

XX The human lupus antigen (La) is diagnostic for Sjogren's syndrome,
 CC as well as occurring in systemic lupus erythematosus patients. The
 CC La protein is clinically related to the Ro protein that is highly
 CC common among autoimmune patients. La and Ro antigens sometimes
 CC reside on the same cellular ribonucleoprotein particle; most La
 CC patients contain some Ro antibodies and vice versa. La cDNA has
 CC been isolated from a human liver library.

XX Sequence 408 AA;

Query Match 90.6%; Score 87; DB 17; Length 408;
 Best Local Similarity 88.9%; Pred. No. 5e-06;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAICQIQIEYFGDF 18

Db 11 AALEAKICHQIEYFQDF 28
|||||:|||||

RESULT 4
ABG08417
ID ABG08417 standard; Protein; 439 AA.

XX AC ABG08417;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #8408.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX XX Dmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX DR N-PSDB; AAS72604.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX PS Claim 20; SEQ ID No 38776; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 439 AA;

Query Match 90.6%; Score 87; DB 22; Length 439;
Best Local Similarity 88.9%; Pred. No. 5.5e-06;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEYFQDF 18

Db 41 AALEAKICHQIEYFQDF 58
|||||:|||||

RESULT 5
AAB58987
ID AAB58987 standard; Protein; 460 AA.
XX AC AAB58987;
XX DT 27-MAR-2001 (first entry)
XX DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 695.
XX DE Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; antitumor; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.

XX OS Homo sapiens.

XX PN WO2000055173-A1.

XX PD 21-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US05881.

XX PR 12-MAR-1999; 99US-0124270.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX DR WPI; 2000-611515/58.

XX DR N-PSDB; AAF21890.

XX PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases

XX PS Claim 11; Page 1149-11150; 1299pp; English.

XX CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antitumor; antitumor; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiac activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemia; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.

XX SQ Sequence 460 AA;

Query Match 90.6%; Score 87; DB 21; Length 460;
Best Local Similarity 88.9%; Pred. No. 5.8e-06;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AALEAKICHQIEYFQDF 18

|||||:|||||

Db 63 AALEAKICHQIEYFGDF 80

RESULT 6

ID ABP41511 standard; Protein; 460 AA.

XX ABP41511;

XX 22-AUG-2002 (first entry)

XX Human ovarian antigen HVAF56, SEQ ID NO:2643.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

XX ovarian cancer; breast cancer; tumour; reproductive system disorder;

XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

XX PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;

XX inflammatory condition; immune disorder; blood disorder;

XX cardiovascular disorder; respiratory disorder; neurological disorder;

XX gastrointestinal disorder; urinary system disorder; drug screening;

XX gene therapy; chromosome mapping; forensic analysis;

XX antibody preparation; cytostatic; immunomodulatory; neuroprotective;

XX antiinflammatory; gynaecological; reproductive.

XX Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US18569.

XX 07-JUN-2000; 2000US-209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

XX N-PSDB; ABQ54588.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,

XX useful in the prevention, treatment and diagnosis of cancer (e.g.

XX ovarian cancer), immune disorders, cardiovascular disorders and

XX neurological diseases -

XX Claim 11; SEQ ID No 2643; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-

XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also

XX encompasses polypeptides 90% identical and polynucleotides 95% identical

XX to the sequences of the invention. The invention additionally relates to

XX recombinant vectors and host cells comprising human ovarian antigen

XX polynucleotides, antibodies against human ovarian antigens, and the use

XX of ovarian antigen polynucleotides and polypeptides in diagnosing,

XX treating, prognosing or preventing various ovary and/or breast-related

XX disorders. Such conditions include ovarian cancer and breast cancer, and

XX metastatic tumours of ovarian or breast origin, reproductive system

XX disorders (e.g., infertility, disorders of pregnancy, anovulation,

XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine

XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic

XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and

XX vaginitis), immune disorders (e.g., congenital and acquired

XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),

XX blood-related disorders (e.g., anaemia), cardiovascular disorders,

XX respiratory disorders, neurological disorders, gastrointestinal disorders

XX and urinary system disorders. Ovarian antigen polypeptides and

XX polynucleotides may also be used in screening for compounds which

XX modulate ovarian antigen expression or activity. The polynucleotides may

XX further be used for gene therapy, chromosome mapping, in the

XX identification of individuals and in forensic analysis, and the

XX polypeptides may be used as food additives or to prepare antibodies

XX useful in disease diagnosis, drug targeting and phenotyping. The present

CC sequence represents a human ovarian antigen of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 460 AA;

Query Match 90.6%; Score 87; DB 23; Length 460;

Best Local Similarity 88.9%; Pred. No. 5.8e-06;

Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEYFGDF 18

DB 63 AALEAKICHQIEYFGDF 80

RESULT 7

ID AAR43394 standard; peptide; 21 AA.

XX AAR43394;

XX 12-MAY-1994 (first entry)

XX La/SSB epitope 17.

XX Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;

XX nuclear ribonucleoprotein; rRNP; Sm B/B'; polypeptide; antigen; D;

XX systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';

XX RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

XX Homo sapiens.

XX WO9321223-A.

XX 28-OCT-1993.

XX 13-APR-1993; 93WO-US03484.

XX 13-APR-1992; 92US-0867819.

XX (OKLA) UNIV OKLAHOMA STATE.

XX Harley JB;

XX WPI; 1993-351658/44.

XX New linear epitope(s) for human auto-antibodies - from the

XX Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used

XX for diagnosing and treating auto-immune disorders e.g. systemic

XX lupus erythematosus

XX Claim 1; Page 30; 43pp; English.

XX The sequences given in AAR43391-562 are linear epitopes which are

XX derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,

XX the 70 kD nuclear ribonucleoprotein (rRNP) and the Sm B/B'

XX polypeptide. These antigens are common in systemic lupus

XX erythematosus (SLE) and closely related disorders. The Ro/SSA family

XX of proteins has been shown to have several molecular forms which are

XX defined by the molecular weight of the antigen identified. The major

XX form has a molecular weight of 60 kD and two additional forms have

XX molecular weights of 52 and 54 kD. La/SSB is also a member of this

XX group of autoantibodies and binds small RNAs with a polynucleotide

XX terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin

XX positive sera. La/SSB has been shown to be a 46-50 kD monomeric

XX phosphoprotein which associates with RNA polymerase III transcripts.

XX Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and

XX U5 RNA. Anti-Sm antibodies may be directed against one or a

XX combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),

XX E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be

XX used for preventing, treating or screening autoimmune disorders,

XX especially SLE or Sjogrens syndrome (SS). They bind to a human

Db 51 ERAIRQVEYYFGD 64

KW Galactokinase; deficiency; galactosemia; diagnosis; therapy;
KW vector; transgenic animal; cataract.

Db 51 ERAIRQVEYYFGD 64

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ALEAICQOIEYF 15
| : : : : : |
Db 164 AARQVCQAHSF 177

RESULT 13
AAB58960
ID AAB58960 standard; Protein; 406 AA.
AC AAB58960;
XX
XX
DT 27-MAR-2001 (first entry)
XX
XX
DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 668.
XX
XX Human: breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.
XX
OS Homo sapiens.
XX
XX WO200055173-A1.
PN
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05881.
XX
XX 12-MAR-1999; 99US-0124270.
PR
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM;
PI
XX WPI: 2000-611515/58.
DR N-PSDB; AAF21863.
XX
XX New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases -
XX
PS Claim 11; Page 1116-1118; 1299pp; English.
XX
CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antitumor; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiant activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemia; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.
XX
SQ Sequence 406 AA;

Query Match 45.8%; Score 44; DB 21; Length 406;

Best Local Similarity 57.1%; Pred. No. 59;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ALEAICQOIEYF 15
| : : : : : |
Db 178 AARQVCQAHSF 191

RESULT 14
AAG47714
ID AAG47714 standard; Protein; 913 AA.
XX
XX
AC AAG47714;
XX
XX
DT 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 60168.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0128785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142058.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144009.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 44.8%; Score 43; DB 21; Length 913;
Best Local Similarity 58.3%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 6 QICQIEYFSD 17
Db 116 KIVNQEYFSD 127

RESULT 15
AAG47713
ID AAG47713 standard; Protein; 923 AA.
XX AC AAG47713;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 60167.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123348.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140895.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 23-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 44.8%; Score 43; DB 21; Length 923;
Best Local Similarity 58.3%; Pred. No. 2.2e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Caps 0;

Qy 6 QICQOIEYFGD 17
: : : : :
Db 126 KIVNQVEYFSD 137

Search completed: April 23, 2003, 13:27:08
Job time : 30.5169 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:34:48 ; Search time 20.0225 Seconds
(without alignments)
185.234 Million cell updates/sec

Title: US-09-836-073-2

Perfect score: 96

Sequence: 1 AALEAQICQIEYVFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 9297

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	31.2	24	Q05616	Q05616 staphylococ
2	29	30.2	16	Q9R963	Q9R963 helicobacte
3	28	29.2	11	Q9UC46	Q9UC46 homo sapien
4	27	28.1	18	Q16028	Q16028 homo sapien
5	27	28.1	25	Q9XJZ7	Q9XJZ7 aureocoumbra
6	25.5	26.6	24	Q38270	Q38270 bacterioph
7	25	26.0	23	Q9UCL8	Q9UCL8 homo sapien
8	25	26.0	24	Q3BM09	Q3BM09 spongilla 1
9	25	26.0	25	Q9TGB8	Q9TGB8 alnus crisp
10	25	26.0	25	Q9TGB7	Q9TGB7 alnus glut
11	25	26.0	25	Q9TGB6	Q9TGB6 alnus marit
12	25	26.0	25	Q9TGB5	Q9TGB5 betula alle
13	25	26.0	25	Q9TGB4	Q9TGB4 betula glan
14	25	26.0	25	Q9TGB3	Q9TGB3 betula papy
15	25	26.0	25	Q9TGB2	Q9TGB2 betula verr
16	25	26.0	25	Q9TGB1	Q9TGB1 betula pube

17	25	26.0	25	8	Q9TGB0	Q9TGB0 corylus ave
18	25	26.0	25	8	Q9TGA9	Q9TGA9 corylus col
19	25	26.0	25	8	Q9TGA8	Q9TGA8 corylus cor
20	25	26.0	25	8	Q9TGA7	Q9TGA7 ostrya virg
21	25	26.0	25	8	Q9TGA6	Q9TGA6 quercus rub
22	25	26.0	25	8	Q9TGB3	Q9TGB3 carpinus ca
23	24.5	25.5	14	11	Q9TGH4	Q9TGH4 mus musculu
24	24	25.0	13	8	Q9THR8	Q9THR8 bryopsis sp
25	24	25.0	13	12	Q9EIV4	Q9EIV4 hepatitis b
26	24	25.0	13	12	Q9EIV3	Q9EIV3 hepatitis b
27	24	25.0	13	12	Q9EIV2	Q9EIV2 hepatitis b
28	24	25.0	13	12	Q9EIV1	Q9EIV1 hepatitis b
29	24	25.0	13	12	Q9EIV0	Q9EIV0 hepatitis b
30	24	25.0	13	12	Q9EIU9	Q9EIU9 hepatitis b
31	24	25.0	13	12	Q9EIU8	Q9EIU8 hepatitis b
32	24	25.0	13	12	Q9EIU6	Q9EIU6 hepatitis b
33	24	25.0	13	12	Q9EIU5	Q9EIU5 hepatitis b
34	24	25.0	13	12	Q9EIU4	Q9EIU4 hepatitis b
35	24	25.0	13	12	Q9EIU3	Q9EIU3 hepatitis b
36	24	25.0	13	12	Q9EIU2	Q9EIU2 hepatitis b
37	24	25.0	13	12	Q9EIU1	Q9EIU1 hepatitis b
38	24	25.0	13	12	Q9EIU0	Q9EIU0 hepatitis b
39	24	25.0	13	12	Q9EIT9	Q9EIT9 hepatitis b
40	24	25.0	13	12	Q9EIT8	Q9EIT8 hepatitis b
41	24	25.0	16	2	Q45663	Q45663 bacillus su
42	24	25.0	20	4	Q9UCER	Q9UCER homo sapien
43	24	25.0	23	4	Q96BE3	Q96BE3 homo sapien
44	23	24.0	20	5	Q9TWR0	Q9TWR0 blattella g
45	23	24.0	20	10	Q9S8X5	Q9S8X5 glycine max

ALIGNMENTS

RESULT 1

Q05616	PRELIMINARY;	PRT;	24 AA.
ID	Q05616		
AC	Q05616;		
DT	01-NOV-1996 (TREMREL. 01, Created)		
DT	01-NOV-1996 (TREMREL. 01, Last sequence update)		
DT	01-NOV-1998 (TREMREL. 08, Last annotation update)		
DE	3-dehydroquinase synthase (EC 4.6.1.3) (Fragment).		
GN	ARO6.		
OS	Staphylococcus aureus.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;		
OC	Staphylococcus.		
OX	NCBI_TaxID=1280;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-8325-4;		
RX	MEDLINE-93381456; PubMed-83711108;		
RA	O'Connell C.M., Pattee P., Foster T.J.;		
RT	"Sequence and mapping of the araA gene of Staphylococcus aureus 8325-4".		
RL	J. Gen. Microbiol. 139:1449-1460(1993).		
CC	- - CATALYTIC ACTIVITY: 7-PHOSPHO-3-DEOXY-ARABINO-HEPTULOSONATE - 3-DEHYDROQUINATE + ORTHOPHOSPHATE.		
CC	- - PATHWAY: SECOND STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).		
CC	- - SUBUNIT: MONOMER (BY SIMILARITY).		
DR	EMBL; L05004; AAA71896.1; -.		
KW	Aromatic amino acid biosynthesis; Lyase.		
FT	NON_TER 1		
SQ	SEQUENCE 24 AA; 2857 MW; 1C4C8D20F7A4856E CRC64;		

Query Match 31.2%; Score 30; DB 2; Length 24;

Best Local Similarity 50.0%; Pred. No. 4.1e-02;

Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 COQIEYVF 15

DB 16 CEQLKTYF 23

```

RESULT 2
Q9R963 ID Q9R963 PRELIMINARY; PRT; 16 AA.
AC Q9R963;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PepC (Fragment).
GN FEPC.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F31;
RX MEDLINE=98453456; PubMed=9780260;
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,
RA Kuriyama M.;
RT "Full-length sequence analysis of the vacA gene from cytotoxic and
RT nontoxic Helicobacter pylori."
RL J. Infect. Dis. 178:1391-1398(1998).
DR EMBL; AF049623; AAD04263.1; -.
FT NON_TER 1
FT NON_TER 16
FT NON_TER 16
SQ SEQUENCE 16 AA; 1938 MW; 3C33D403AEE61428 CRC64;

Query Match 30.2%; Score 29; DB 2; Length 16;
Best Local Similarity 55.6%; Pred. No. 3.9e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 10 QIEYFGDF 18
I:|:|:|
Db 7 QVEYAFNF 15

RESULT 3
Q9UC46 ID Q9UC46 PRELIMINARY; PRT; 11 AA.
AC Q9UC46;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Neutrophil inhibitor peptide, NIP-POLYMORPHONUCLEAR neutrophil
DE inhibitor peptide.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=96326114; PubMed=8703476;
RA Cooper J.A.Jr., Culbreth R.R.;
RT "Characterization of a neutrophil inhibitor peptide harvested from
RT human bronchial lavage; homology to influenza A nucleoprotein."
RL Am. J. Respir. Cell Mol. Biol. 15:207-215(1996).
SQ SEQUENCE 11 AA; 1262 MW; 951A1C3279C9DB45 CRC64;

Query Match 29.2%; Score 28; DB 4; Length 11;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 13 YFFGD 17
I:|:|:|
Db 5 YFFGD 9

RESULT 4
Q16028 ID Q16028 PRELIMINARY; PRT; 18 AA.
AC Q16028;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

```

```

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE OCR1-1 protein (Fragment).
GN OCR1-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93278398; PubMed=8504307;
RA Leachey A.M., Charnas L.R., Nussbaum R.L.;
RT "Nonsense mutations in the OCR1-1 gene in patients with the
RT oculocerebrorenal syndrome of Lowe."
RL Hum. Mol. Genet. 2:461-463(1993).
DR EMBL; S62084; AAD13933.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 18 AA; 2055 MW; 61356C825F7AF33D CRC64;

Query Match 28.1%; Score 27; DB 4; Length 18;
Best Local Similarity 36.4%; Pred. No. 9.7e+02;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AALEAQICQOI 11
I:|:|:|
Db 1 SAYDPRICRQL 11

RESULT 5
Q9XJZ7 ID Q9XJZ7 PRELIMINARY; PRT; 25 AA.
AC Q9XJZ7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit
DE (Fragment).
GN RECS.
OS Aureocoumbra lagunensis.
OG Chloroplast.
OC Eukaryota; stramenopiles; Pelagophyceae; Aureocoumbra.
OX NCBI_TaxID=44058;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCMP1681, CCMP1509, AND CCMP1510;
RA Bailey J.C., Andersen R.A.;
RT "Analysis of clonal cultures of the brown tide algae Aureococcus and
RT Aureocoumbra (Pelagophyceae) using 18S rRNA, rbcL and rubisco spacer
RT sequences."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF117782; AAD39106.1; -.
DR EMBL; AF117780; AAD39104.1; -.
DR EMBL; AF117781; AAD39105.1; -.
DR InterPro; IPR000894; RUBISCO_small.
DR Pfam; PF00101; RUBISCO_small; 1.
KW Chloroplast.
FT NON_TER 25
FT NON_TER 25
SQ SEQUENCE 25 AA; 2972 MW; 61A96BD83DB94118 CRC64;

Query Match 28.1%; Score 27; DB 8; Length 25;
Best Local Similarity 62.5%; Pred. No. 1.4e+03;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 QICQIQEY 13
I:|:|:|
Db 18 QIRKQVEY 25

RESULT 6
Q38270 ID Q38270 PRELIMINARY; PRT; 24 AA.
AC Q38270;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

```


OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OX eurosids I; Fagales; Betulaceae; Alnus.
 RN NCBI_TaxID=3517;
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99261656; PubMed=10331271;
 RA Laroche J., Bousquet J.;
 RT "Evolution of the mitochondrial rps3 intron in perennial and annual
 angiosperms and homology to nad5 intron 1.";
 RL Mol. Biol. Evol. 16:441-452(1999).
 DR EMBL; AF080076; AAD50063.1; -;
 DR InterPro; IPR001351; Ribosomal_S3.
 DR Pfam; PF00417; Ribosomal_S3_N; 1.
 KW Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;
 Query Match 26.0%; Score 25; DB 8; Length 25;
 Best Local Similarity 60.0%; Pred. No. 2.9e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 12 EYFG 16
 Db :||:|
 4 DYYG 8

RESULT 11
 Q9TGB6
 ID Q9TGB6 PRELIMINARY; PRT; 25 AA.
 AC Q9TGB6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Ribosomal protein small 3 (Fragment).
 GN RPS3.
 OS Alnus maritima.
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OX eurosids I; Fagales; Betulaceae; Alnus.
 RN NCBI_TaxID=21015;
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99261656; PubMed=10331271;
 RA Laroche J., Bousquet J.;
 RT "Evolution of the mitochondrial rps3 intron in perennial and annual
 angiosperms and homology to nad5 intron 1.";
 RL Mol. Biol. Evol. 16:441-452(1999).
 DR EMBL; AF080077; AAD50064.1; -;
 DR InterPro; IPR001351; Ribosomal_S3.
 DR Pfam; PF00417; Ribosomal_S3_N; 1.
 KW Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;
 Query Match 26.0%; Score 25; DB 8; Length 25;
 Best Local Similarity 60.0%; Pred. No. 2.9e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 12 EYFG 16
 Db :||:|
 4 DYYG 8

RESULT 12
 Q9TGB5
 ID Q9TGB5 PRELIMINARY; PRT; 25 AA.
 AC Q9TGB5;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Ribosomal protein small 3 (Fragment).
 GN RPS3.
 OS Betula alleghaniensis.
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OX eurosids I; Fagales; Betulaceae; Betula.
 RN NCBI_TaxID=21017;
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99261656; PubMed=10331271;
 RA Laroche J., Bousquet J.;
 RT "Evolution of the mitochondrial rps3 intron in perennial and annual
 angiosperms and homology to nad5 intron 1.";
 RL Mol. Biol. Evol. 16:441-452(1999).
 DR EMBL; AF080078; AAD50065.1; -;
 DR InterPro; IPR001351; Ribosomal_S3.
 DR Pfam; PF00417; Ribosomal_S3_N; 1.
 KW Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;
 Query Match 26.0%; Score 25; DB 8; Length 25;
 Best Local Similarity 60.0%; Pred. No. 2.9e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 12 EYFG 16
 Db :||:|
 4 DYYG 8

RESULT 13
 Q9TGB4
 ID Q9TGB4 PRELIMINARY; PRT; 25 AA.
 AC Q9TGB4;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Ribosomal protein small 3 (Fragment).
 GN RPS3.
 OS Betula glandulosa.
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OX eurosids I; Fagales; Betulaceae; Betula.
 RN NCBI_TaxID=21018;
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99261656; PubMed=10331271;
 RA Laroche J., Bousquet J.;
 RT "Evolution of the mitochondrial rps3 intron in perennial and annual
 angiosperms and homology to nad5 intron 1.";
 RL Mol. Biol. Evol. 16:441-452(1999).
 DR EMBL; AF080079; AAD50066.1; -;
 DR InterPro; IPR001351; Ribosomal_S3.
 DR Pfam; PF00417; Ribosomal_S3_N; 1.
 KW Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;
 Query Match 26.0%; Score 25; DB 8; Length 25;
 Best Local Similarity 60.0%; Pred. No. 2.9e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 12 EYFG 16
 Db :||:|
 4 DYYG 8

RESULT 14
 Q9TGB3

ID Q9TGB3 PRELIMINARY; PRT; 25 AA.
 AC Q9TGB3;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Ribosomal protein small 3 (Fragment).
 GN RPS3.
 OS Betula papyrifera (Paper birch).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fagales; Betulaceae; Betula.
 OX NCBI_TaxID=3507;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99261656; PubMed=10331271;
 RA Laroche J., Bousquet J.;
 RT "Evolution of the mitochondrial rps3 intron in perennial and annual
 angiosperms and homology to nad5 intron 1.";
 RL Mol. Biol. Evol. 16:441-452(1999).
 DR EMBL; AF080080; AAD50067.1; -;
 DR InterPro; IPR001351; Ribosomal_S3.
 DR Pfam; PF00417; Ribosomal_S3_N; 1.
 KW Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 26.0%; Score 25; DB 8; Length 25;
 Best Local Similarity 60.0%; Pred. No. 2.9e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
 Db :||:|
 Db 4 DYYG 8

RESULT 15

Q9TGB2
 ID Q9TGB2 PRELIMINARY; PRT; 25 AA.
 AC Q9TGB2;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Ribosomal protein small 3 (Fragment).
 GN RPS3.
 OS Betula verrucosa (White birch) (Betula pendula).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fagales; Betulaceae; Betula.
 OX NCBI_TaxID=3505;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99261656; PubMed=10331271;
 RA Laroche J., Bousquet J.;
 RT "Evolution of the mitochondrial rps3 intron in perennial and annual
 angiosperms and homology to nad5 intron 1.";
 RL Mol. Biol. Evol. 16:441-452(1999).
 DR EMBL; AF080081; AAD50068.1; -;
 DR InterPro; IPR001351; Ribosomal_S3.
 DR Pfam; PF00417; Ribosomal_S3_N; 1.
 KW Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 26.0%; Score 25; DB 8; Length 25;
 Best Local Similarity 60.0%; Pred. No. 2.9e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
 Db :||:|

Db 4 DYYG 8

Search completed: April 23, 2003, 13:47:13
 Job time : 21.0225 secs

CC autoantibody and may therefore be used as vaccines.

XX Sequence 21 AA;

Query Match 71.6%; Score 73; DB 14; Length 21;

Best Local Similarity 100.0%; Pred. No. 4.7e-05;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICHQIEYFGDF 18

Db 1 ICHQIEYFGDF 12

RESULT 8

ABB65316
ID ABB65316 standard; Protein; 390 AA.

XX AC ABB65316;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 22740.

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-151637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL09419.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -

XX PS Disclosure; SEQ ID NO 22740; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins
XX CC (ABB57737-ABB72072).

XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 390 AA;

Query Match 55.9%; Score 57; DB 22; Length 390;

Best Local Similarity 62.5%; Pred. No. 0.46;

Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QOEAKICHQIEYFGD 17

Db 49 KOEAKIRQVEYFGD 64

RESULT 9

AAG47714

ID AAG47714 standard; Protein; 913 AA.

XX AC AAG47714;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 60168.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 23-APR-1999; 99US-0130891.

XX PR 28-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 04-MAY-1999; 99US-0132484.

XX PR 05-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 07-MAY-1999; 99US-0132863.

XX PR 11-MAY-1999; 99US-0134256.

XX PR 14-MAY-1999; 99US-0134218.

XX PR 14-MAY-1999; 99US-0134219.

XX PR 14-MAY-1999; 99US-0134221.

XX PR 14-MAY-1999; 99US-0134370.

XX PR 18-MAY-1999; 99US-0134768.

XX PR 19-MAY-1999; 99US-0134941.

XX PR 20-MAY-1999; 99US-0135124.

XX PR 21-MAY-1999; 99US-0135353.

XX PR 24-MAY-1999; 99US-0135629.

XX PR 25-MAY-1999; 99US-0136021.

XX PR 27-MAY-1999; 99US-0136392.

XX PR 28-MAY-1999; 99US-0136782.

XX PR 01-JUN-1999; 99US-0137222.

XX PR 03-JUN-1999; 99US-0137528.

XX PR 04-JUN-1999; 99US-0137502.

XX PR 07-JUN-1999; 99US-0137724.

XX PR 08-JUN-1999; 99US-0138094.

XX PR 10-JUN-1999; 99US-0138540.

XX PR 10-JUN-1999; 99US-0138847.

XX PR 14-JUN-1999; 99US-0139119.

XX PR 16-JUN-1999; 99US-0139452.

XX PR 16-JUN-1999; 99US-0139453.

XX PR 17-JUN-1999; 99US-0139492.

XX PR 18-JUN-1999; 99US-0139454.

XX PR 18-JUN-1999; 99US-0139455.

XX PR 18-JUN-1999; 99US-0139456.

XX PR 18-JUN-1999; 99US-0139457.

XX PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145313.
 PR 27-JUL-1999; 99US-0145318.
 PR 27-JUL-1999; 99US-0145319.
 PR 28-JUL-1999; 99US-0145351.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 47.1%; Score 48; DB 21; Length 913;
 Best Local Similarity 66.7%; Pred. NO. 34;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIEYFGD 17
 || :||| |
 Db 116 KIVNQVEYFSD 127

RESULT 10.
 AAG47713
 ID AAG47713 standard; Protein; 923 AA.
 XX
 AC AAG47713;
 XX
 DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 60167.
DE Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
KW
XX Arabidopsis thaliana.
OS
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 28-MAR-1999; 99US-0126284.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134236.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137744.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.
XX 02-JUL-1999; 99US-0142055.
XX 06-JUL-1999; 99US-0142390.
XX 08-JUL-1999; 99US-0142803.
XX 09-JUL-1999; 99US-0142920.
XX 12-JUL-1999; 99US-0142977.
XX 13-JUL-1999; 99US-0143542.
XX 14-JUL-1999; 99US-0143624.
XX 15-JUL-1999; 99US-0144005.
XX 16-JUL-1999; 99US-0144085.
XX 16-JUL-1999; 99US-0144086.
XX 19-JUL-1999; 99US-0144325.
XX 19-JUL-1999; 99US-0144331.
XX 19-JUL-1999; 99US-0144332.
XX 19-JUL-1999; 99US-0144333.
XX 19-JUL-1999; 99US-0144334.
XX 19-JUL-1999; 99US-0144335.
XX 20-JUL-1999; 99US-0144352.
XX 20-JUL-1999; 99US-0144632.
XX 20-JUL-1999; 99US-0144884.
XX 21-JUL-1999; 99US-0144814.
XX 21-JUL-1999; 99US-0145086.
XX 21-JUL-1999; 99US-0145088.
XX 22-JUL-1999; 99US-0145085.
XX 22-JUL-1999; 99US-0145087.
XX 22-JUL-1999; 99US-0145089.
XX 22-JUL-1999; 99US-0145192.
XX 23-JUL-1999; 99US-0145145.
XX 23-JUL-1999; 99US-0145218.
XX 23-JUL-1999; 99US-0145224.
XX 26-JUL-1999; 99US-0145276.
XX 27-JUL-1999; 99US-0145913.
XX 27-JUL-1999; 99US-0145918.
XX 27-JUL-1999; 99US-0145919.
XX 28-JUL-1999; 99US-0145951.
XX 02-AUG-1999; 99US-0146386.
XX 02-AUG-1999; 99US-0146388.
XX 02-AUG-1999; 99US-0146389.
XX 03-AUG-1999; 99US-0147038.
XX 04-AUG-1999; 99US-0147204.
XX 04-AUG-1999; 99US-0147302.
XX 05-AUG-1999; 99US-0147192.
XX 05-AUG-1999; 99US-0147260.
XX 06-AUG-1999; 99US-0147303.
XX 06-AUG-1999; 99US-0147416.
XX 09-AUG-1999; 99US-0147493.
XX 10-AUG-1999; 99US-0147935.
XX 10-AUG-1999; 99US-0148171.
XX 11-AUG-1999; 99US-0148319.
XX 12-AUG-1999; 99US-0148341.
XX 13-AUG-1999; 99US-0148565.
XX 16-AUG-1999; 99US-0148684.
XX 17-AUG-1999; 99US-0149175.
XX 18-AUG-1999; 99US-0149426.
XX 20-AUG-1999; 99US-0149722.
XX 20-AUG-1999; 99US-0149723.
XX 20-AUG-1999; 99US-0149929.
XX 23-AUG-1999; 99US-0149930.
XX 23-AUG-1999; 99US-0149930.
XX 25-AUG-1999; 99US-0150566.
XX 26-AUG-1999; 99US-0150884.
XX 27-AUG-1999; 99US-0151065.
XX 27-AUG-1999; 99US-0151066.

```

PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159384.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 47.1%; Score 48; DB 21; Length 923;
Best Local Similarity 66.7%; Pred. No. 34;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIEYFSD 17
   ||:|||||
Db 126 KIVNQVEYFSD 137

RESULT 11
AAG47712
ID AAG47712 standard; Protein; 993 AA.
AC AAG47712;
XX
XX
DT 18-OCT-2000 (first entry)
DE
DE Arabidopsis thaliana protein fragment SEQ ID NO: 60166.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX

```

```

OS Arabidopsis thaliana.
XX EF1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.

```

PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 03-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 26-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 28-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 23-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 26-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 47.1%; Score 48; DB 21; Length 993;
 Best Local Similarity 66.7%; Pred. No. 37;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIYYFGD 17
 || :||:|||||
 Db 196 KIVNQVEYFSD 207

RESULT 12
 ABB59107
 ID ABB59107 standard; Protein; 1900 AA.

XX ABB59107;
 AC
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 4113.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX

PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL03210.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 4113; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1900 AA;
Query Match 44.1%; Score 45; DB 22; Length 1900;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Qy 1 QQQEKICHQIEY 14
Db 1275 QQRELDCHREFF 1288
RESULT 13
ABG27058
ID ABG27058 standard; Protein; 135 AA.
XX
AC ABG27058;
XX
DT 18-FEB-2002 (first entry)
XX
DE Human human diagnostic protein #27049.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS91245.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity -
XX
PS Claim 20; SEQ ID NO 57417; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 135 AA;
Query Match 43.1%; Score 44; DB 22; Length 135;
Best Local Similarity 70.0%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 4 EAKICHQIEY 13
Db 9 EKRICHQIEF 18
RESULT 14
ABB89645
ID ABB89645 standard; Protein; 224 AA.
XX
AC ABB89645;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 2021.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US16450.
XX
PR 19-MAY-2000; 2000US-205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR N-PSDB; ABL90054.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative

PT disorders -

PS Claim 11; SEQ ID NO 2021; 2081pp + Sequence Listing; English.

The invention relates to novel genes (ABL99449-ABL90853), and proteins (ABB9040-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anaemia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct_sequences.

CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 224 AA;

Query Match 43.1%; Score 44; DB 23; Length 224;
Best Local Similarity 70.0%; Pred. NO. 33;
Matches 7; Conservative 2; Mismatches 1; Indels

OV 4 EAKICHOIEY 13

112 EKRRCHQIEF 121

RESULT 15

RECOIL IS
ABB59213
ID ABB59213 standard: Protein: 277 AA.

AC ABB59213:

26-MAR-2002 (first entry)

DE Prosopihila melanoqaster polypeptide SEO ID NO 4431.

AA Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical
KW
KW

05 *Drosophila melanogaster*.

XX PN WO200171042-A2

27-SEP-2001.

23-MAR-2001: 2001WO-US09231.

23-MAR-2000: 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI .Venter JC, Adams M, Li PWD, Myers EW:

WPI: 2001-656860/75.

DR N-PSDB; ABL03316

PT. New isolated nucleic acid detection reagent for detecting 1000 or more
PT. genes from *Drosophila* and for elucidating cell signalling and cell-cell
PT. interactions -

PS Disclosure; SEQ ID NO 4431; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from *Drosophila*. The invention is capable of detecting 1000 or more isolated nucleic acid detection reagent

useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct_sequences.

CC (AB557/31=AB5720/Z).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

AA	Sequence	277 AA:
SO		

Query Match 43.18: Score 44: DB 22: Length 277;

Best Local Similarity 50.0%; Pred. No. 42;

Matches	7;	Conservative	3;	Mismatches	4;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	----	--------	----	------	----

QY 1 QQQEAKICHQIEYY 14

Search completed: April 23, 2003, 13:27:11
Job time : 29.5169 secs

Result No.	Score	Query Match	Length	DB ID	Description
1	33	32.4	24	5	Q9BM09
2	31	30.4	24	3	Q9BM09 spongilla l
3	31	30.4	24	3	Q07140 saccharomyc
3	29	28.4	16	4	Q9R963 helicobacte
4	28	27.5	11	4	Q9UC46 homo sapien
5	28	27.5	24	2	Q05616 staphylococ
6	27	26.5	20	10	Q9UF03
7	27	26.5	23	4	Q9UCL8
7	26	25.5	8	4	Q9BY15
8	26	25.5	8	6	Q9BFA0
9	26	25.5	8	6	Q9BFA0 macaca mula
10	26	25.5	8	6	Q9BF99
11	26	25.5	24	7	Q9TNS7
12	26	25.5	24	7	Q9TNS6
13	25	24.5	25	8	Q9TGB8
14	25	24.5	25	8	Q9TGB7
15	25	24.5	25	8	Q9TGB6
16	25	24.5	25	8	Q9TGB5

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Class II transposable element Tyl-17 5' end (Fragment).
 GN YCLO19W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MD 40-4C;
 RX MEDLINE=85242090; PubMed=2989787;
 RA Fulton A.M., Mellor J., Dobson M.J., Chester J., Warrington J.R.,
 RA Indge K.J., Oliver S.G., de la Paz P., Wilson W., Kingsman A.J.,
 RA Kingsman S.M.;
 RA "Variants within the yeast Ty sequence family encode a class of
 RT structurally conserved proteins.";
 RL Nucleic Acids Res. 13:4097-4111(1985).
 DR EMBL; X02546; CAA26399.1; -;
 DR SGD; S0000524; YCLO19W.
 FT NON_TER 24
 SQ SEQUENCE 24 AA; 2642 MW; A5B0BDF527A58B3 CRC64;
 Query Match 30.4%; Score 31; DB 3; Length 24;
 Best Local Similarity 50.0%; Pred. No. 3.1e+02;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QQQEAKICHOIE 12
 Db 11 QQKESKPTHTID 22
 RESULT 3
 Q9R963 ID Q9R963 PRELIMINARY; PRT; 16 AA.
 AC Q9R963;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE FepC (Fragment).
 GN FEP.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=F31;
 RX MEDLINE=98453456; PubMed=9780260;
 RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,
 RA Kuriyama M.;
 RA "Full-length sequence analysis of the vacA gene from cytotoxic and
 RT noncytotoxic Helicobacter pylori.";
 RL J. Infect. Dis. 178:1391-1398(1998).
 DR EMBL; AF049623; AAD04263.1; -;
 DR NON_TER 1
 FT NON_TER 16
 SQ SEQUENCE 16 AA; 1938 MW; 3C33D403AEE51428 CRC64;
 Query Match 28.4%; Score 29; DB 2; Length 16;
 Best Local Similarity 55.6%; Pred. No. 4.4e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 10 QIEYFGDF 18
 Db 7 QVEYAFNF 15
 RESULT 4
 Q9UC46 ID Q9UC46 PRELIMINARY; PRT; 11 AA.
 AC Q9UC46;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE Neutrophil inhibitor peptide, NIP-POLYMORPHONUCLEAR neutrophil
 DE inhibitor peptide.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96326114; PubMed=8703476;
 RA Cooper J.A.Jr., Culbreth R.R.;
 RA "Characterization of a neutrophil inhibitor peptide harvested from
 RT human bronchial lavage: homology to influenza A nucleoprotein.";
 RL Am. J. Respir. Cell Mol. Biol. 15:207-215(1996).
 SQ SEQUENCE 11 AA; 1262 MW; 951AIC3279C9DB45 CRC64;
 Query Match 27.5%; Score 28; DB 4; Length 11;
 Best Local Similarity 80.0%; Pred. No. 4.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 13 YFEGD 17
 Db 5 YFEGD 9
 RESULT 5
 Q05616 ID Q05616 PRELIMINARY; PRT; 24 AA.
 AC Q05616;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE 3-dehydroquinase synthase (EC 4.6.1.3) (Fragment).
 GN AROB.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8325-4;
 RX MEDLINE=93381456; PubMed=8371108;
 RA O'Connell C.M., Pattee P., Foster T.J.;
 RT "Sequence and mapping of the araG gene of Staphylococcus aureus 8325-4.";
 RL J. Gen. Microbiol. 139:1449-1460(1993).
 CC -1- CATALYTIC ACTIVITY: 7-PHOSPHO-3-DEOXY-ARABINO-HEPTULOSONATE - 3-
 CC DEHYDROQUINATE + ORTHOPHOSPHATE.
 CC -1- PATHWAY: SECOND STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE
 CC AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC EMBL; L05004; AAY1896.1; -;
 DR Aromatic amino acid biosynthesis; Lyase.
 FT NON_TER 1
 SQ SEQUENCE 24 AA; 2857 MW; 1C4C8D20F7A4856E CRC64;
 Query Match 27.5%; Score 28; DB 2; Length 24;
 Best Local Similarity 50.0%; Pred. No. 9.7e+02;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 8 CHQIEYF 15
 Db 16 CEQLKTYF 23
 RESULT 6
 Q9FUY3 ID Q9FUY3 PRELIMINARY; PRT; 20 AA.
 AC Q9FUY3;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```

DE MADS-box protein (Fragment).
GN TKZ-1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. INBRED A189; TISSUE=LEAF;
RA Ermishev V.V., Zbrodina M.V., Karyagina A.S., Naroditsky B.S.,
RA Khavkin E.E.;
RT "An unusual K-box sequence of a maize MADS-box gene.";
RL Russ. J. Plant Physiol. 47:555-557(2000).
DR EMBL: AF276682; AAG10438.1; -.
FT NON_TER 1
FT NON_TER 20
FT NON_TER 20
SQ SEQUENCE 20 AA; 2477 MW; 0DFD32A46903C4D0 CRC64;

Query Match 26.5%; Score 27; DB 10; Length 20;
Best Local Similarity 41.7%; Pred. No. 1.2e+03;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QQEAQKICHQIE 12
   ||:|:|:|
Db 2 QQESPXLRNQIQ 13

RESULT 7
Q9UC18 PRELIMINARY; PRT; 23 AA.
ID Q9UC18
AC Q9UC18
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE GLYCOSAMINOGLYCAN-BEARING POLYPEPTIDE (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE
RX MEDLINE=93098810; PubMed=1463459;
RA Bonnet F., Perin J.P., Maillet P., Jolles P., Alliel P.M.;
RT "Characterization of a human seminal plasma glycosaminoglycan-bearing
RT polypeptide.";
RL Biochem. J. 288:565-569(1992).
SQ SEQUENCE 23 AA; 2334 MW; E469397E6B2EA794 CRC64;

Query Match 26.5%; Score 27; DB 4; Length 23;
Best Local Similarity 35.3%; Pred. No. 1.4e+03;
Matches 6; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 QQEAQKICHQIEYFGDF 18
   ||:|:|:|
Db 1 KQAVSCSEEQETXGDF 17

RESULT 8
Q9BY15 PRELIMINARY; PRT; 8 AA.
ID Q9BY15
AC Q9BY15
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE cAMP responsive element moderator (Fragment).
GN CREM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL: AY011664; AAG47575.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 8 AA; 1006 MW; DF02C331EEAB572A CRC64;

Query Match 25.5%; Score 26; DB 4; Length 8;
Best Local Similarity 60.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQIE 12
   ||:|:|
Db 4 CHKVE 8

RESULT 9
Q9BFA0 PRELIMINARY; PRT; 8 AA.
ID Q9BFA0
AC Q9BFA0
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE cAMP responsive element moderator (Fragment).
GN CREM.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL: AY011661; AAG47572.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 8 AA; 1006 MW; DF02C331EEAB572A CRC64;

Query Match 25.5%; Score 26; DB 6; Length 8;
Best Local Similarity 60.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQIE 12
   ||:|:|
Db 4 CHKVE 8

RESULT 10
Q9BF99 PRELIMINARY; PRT; 8 AA.
ID Q9BF99
AC Q9BF99
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE cAMP responsive element moderator (Fragment).
GN CREM.
OS Hylobates concolor (crested gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=29089;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL: AY011662; AAG47573.1; -.

```

FT NON_TER 1 1006 MW; DF02C331EBAB572A CRC64;
SQ SEQUENCE 8 AA; 25.5%; Score 26; DB 6; Length 8;
Query Match Best Local Similarity 60.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHOIE 12
Db 4 CHKE 8

RESULT 11
Q9TNS7 PRELIMINARY; PRT; 24 AA.
AC Q9TNS7; MEDLINE-93107319; PubMed-1469092;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE Human leucocyte antigen beta chain DR molecule HLA-DRB1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93107319; PubMed-1469092;
RA Weyand C.M., Hicok K.C., Hunder G.G., Goronzy J.J.;
RT "The HLA-DRB1 locus as a genetic component in giant cell arteritis."
RT Mapping of a disease-linked sequence motif to the antigen binding site
of the HLA-DR molecule."
RL J. Clin. Invest. 90:2355-2361(1992).
KW MHC.
SQ SEQUENCE 24 AA; 3042 MW; 46FF753670C7A760 CRC64;

Query Match 25.5%; Score 26; DB 7; Length 24;
Best Local Similarity 20.0%; Pred. No. 2.1e+03;
Matches 2; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIEYF 15
Db 3 QVKEFDYF 12

RESULT 12
Q9TNS6 PRELIMINARY; PRT; 24 AA.
AC Q9TNS6; MEDLINE-93107319; PubMed-1469092;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE Human leucocyte antigen beta chain DR molecule HLA-DRB1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93107319; PubMed-1469092;
RA Weyand C.M., Hicok K.C., Hunder G.G., Goronzy J.J.;
RT "The HLA-DRB1 locus as a genetic component in giant cell arteritis."
RT Mapping of a disease-linked sequence motif to the antigen binding site
of the HLA-DR molecule."
RL J. Clin. Invest. 90:2355-2361(1992).
KW MHC.
SQ SEQUENCE 24 AA; 3112 MW; 5C4F753667F7A760 CRC64;

Query Match 25.5%; Score 26; DB 7; Length 24;
Best Local Similarity 20.0%; Pred. No. 2.1e+03;
Matches 2; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIEYF 15
Db 3 QVKEFDYF 12

Db 3 QVKEFDYF 12

RESULT 13
Q9TGB8 PRELIMINARY; PRT; 25 AA.
AC Q9TGB8; MEDLINE-99261656; PubMed-10331271;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Alnus crispa.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=3518;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99261656; PubMed-10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
angiosperms and homology to nads intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080075; AAD50062.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 24.5%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
Db 4 DYYG 8

RESULT 14
Q9TGB7 PRELIMINARY; PRT; 25 AA.
AC Q9TGB7; MEDLINE-99261656; PubMed-10331271;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Alnus glutinosa (Alder).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=3517;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99261656; PubMed-10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
angiosperms and homology to nads intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080076; AAD50063.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 24.5%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
Db 4 DYYG 8

RESULT 14
Q9TGB7 PRELIMINARY; PRT; 25 AA.
AC Q9TGB7; MEDLINE-99261656; PubMed-10331271;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Alnus glutinosa (Alder).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=3517;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99261656; PubMed-10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
angiosperms and homology to nads intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080076; AAD50063.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 24.5%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
Db 4 DYYG 8

RESULT 14
Q9TGB7 PRELIMINARY; PRT; 25 AA.
AC Q9TGB7; MEDLINE-99261656; PubMed-10331271;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Alnus glutinosa (Alder).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=3517;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99261656; PubMed-10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
angiosperms and homology to nads intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080076; AAD50063.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 24.5%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
Db 4 DYYG 8

Best Local Similarity 60.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
:|:|:
Db 4 DYYG 8

RESULT 15

Q9TGB6 PRELIMINARY; PRT; 25 AA.
AC Q9TGB6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Alnus maritima.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=21015;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RT angiosperms and homology to nad5 intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080077; AAD50064.1; -;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF0417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 24.5%; Score 25; DB 8; Length 25;

Best Local Similarity 60.0%; Pred. No. 3.2e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
:|:|:
Db 4 DYYG 8

Search completed: April 23, 2003, 13:47:15
Job time : 21.0225 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:33:13 ; Search time 4.65169 Seconds
(without alignments)
160.495 Million cell updates/sec

Title: US-09-836-073-4

Perfect score: 102
Sequence: 1 QQQEAKICHOIEYFFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1520

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	26.5	19	1 HBB2_UROHA	P18992 uromastix h
2	24	23.5	11	1 CXL1_CONMR	P58807 conus marmo
3	24	23.5	11	1 CXL4_CONMR	P58810 conus marmo
4	24	23.5	16	1 MLB_SQUAC	P01207 squallus aca
5	24	23.5	18	1 MLB_SCYCA	P01206 scyllorhinu
6	24	23.5	25	1 CXOB_CONMA	P05485 conus magus
7	22	21.6	21	1 BFX_ATRBI	P80163 atractaspis
8	22	21.6	21	1 SRPD_ATRBN	P13211 atractaspis
9	22	21.6	25	1 ANDT_ANDAU	P56684 androctonus
10	21	20.6	10	1 GON2_CHEPR	P80678 chelyosoma
11	21	20.6	12	1 TIN2_HOPTI	P82652 hoplobatrach
12	21	20.6	17	1 TPIS_PINPS	P81666 pinus pinas
13	21	20.6	21	1 REV_HV2D2	P15830 human immun
14	21	20.6	22	1 D20_CAEBR	P51558 caenorhabdi
15	21	20.6	22	1 FUC1_RAT	P80347 rattus norv
16	21	20.6	22	1 FUC2_RAT	P80348 rattus norv
17	21	20.6	23	1 NUO5_SOLTU	P80262 solanum tub
18	21	20.6	23	1 PRO3_DACGL	P18690 dactylis gl
19	21	20.6	25	1 SNBP_RAT	P80968 rattus norv
20	20	19.6	10	1 TRNB_RANRI	P29135 rana ridibu
21	20	19.6	14	1 CXAI_CONCN	P81096 brassica na
22	20	19.6	15	1 PC20_BRANA	P56973 conus conso
23	20	19.6	16	1 CXA2_CONMA	P56636 conus magus
24	20	19.6	18	1 OBP_LYMDI	P34173 lymantria d
25	20	19.6	18	1 SFAN_HELAN	P81098 hellanthus
26	20	19.6	20	1 BIP_PHAVU	P80089 phaseolus v
27	20	19.6	20	1 COG4_CHIOP	P34156 chionocete
28	20	19.6	20	1 FIBB_FELCA	P14469 felis silve
29	20	19.6	20	1 TL22_SPIOL	P82796 spinacia ol
30	20	19.6	20	1 YQAH_KLEAE	P56506 klebsiella
31	20	19.6	21	1 NDK_CANAL	Q9UR66 candida alb
32	20	19.6	22	1 LPI_TRIWA	P24335 trimeresuru
33	20	19.6	22	1 LP2_TRIWA	P58930 trimeresuru

34 20 19.6 23 1 XYCL_ACIGB P46365 acinetobact
35 19.5 19.1 20 1 COXN_THUOB P80980 thunnus obe
36 19.5 19.1 23 1 UDP_LACCA P19662 lactobacill
37 19 18.6 8 1 AL17_CARMA P81820 carcinus ma
38 19 18.6 9 1 AL11_CARMA P81814 carcinus ma
39 19 18.6 10 1 TKNB_ONCMY P28500 oncorhynchu
40 19 18.6 11 1 TIN4_HOPTI P82654 hoplobatrach
41 19 18.6 12 1 NO40_SESRO O24369 sesbania ro
42 19 18.6 12 1 TIN3_HOPTI P82653 hoplobatrach
43 19 18.6 14 1 ADF_TENMO P82965 tenebrio mo
44 19 18.6 19 1 FIBB_VULVU P14482 vulpes vulp
45 19 18.6 20 1 C52L_STRTR P81621 streptococc

ALIGNMENTS

RESULT 1
HBB2_UROHA
ID HBB2_UROHA STANDARD; PRT; 19 AA.
AC P18992;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hemoglobin beta-2 chain (fragment).
OS Uromastix hardwickii (Indian spiny-tailed lizard).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Uromastycinae;
OC Uromastix.
OX NCBI_TaxID=40250;
RN [1]
RP SEQUENCE.
RX MEDLINE=84029159; PubMed=6628672;
RA Nagvi S., Zaidi Z.H., von Bahr-Lindstroem H., Carlquist M.,
RA Joernvall H.;
RT "Characterization of hemoglobin from the lizard Uromastix
hardwickii.";
RL FEBS Lett. 162:290-295(1983).
CC -!- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
VARIOUS PERIPHERAL TISSUES.
CC -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -!- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -!- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR; A05305; A05305.
DR InterPro; IPR000971; Globin.
DR PROSITE; PS01033; GLOBIN; PARTIAL.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT NON_TER 1
FT NON_TER 19
SQ SEQUENCE 19 AA; 1914 MW; C40AD8EA30019057 CRC64;

Query Match 26.5%; Score 27; DB 1; Length 19;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 YFGDF 18
Db 1 FFGDF 5

RESULT 2
CXL1_CONMR
ID CXL1_CONMR STANDARD; PRT; 11 AA.
AC P58807;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda-conotoxin CmrVIA.
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]

RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
 RC TISSUE-Venom;
 RA MEDLINE=20564325; PubMed=10988292;
 RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
 RA Seow K.T., Bay B.-H.;
 RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
 RT pattern and protein folding. Isolation and characterization from the
 RT venom of Conus marmoreus";
 RL J. Biol. Chem. 275:39516-39522(2000).
 CC -1- FUNCTION: Inhibits the neuronal noradrenaline transporter.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -1- MASS SPECTROMETRY: MW=1237.93; MW_ERR=0.21; METHOD=Electrospray.
 CC -1- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
 KW Neurotoxin; Toxin; Hydroxylation.
 FT DISULFID 2 11
 FT DISULFID 3 8
 FT MOD_RES 10 10 HYDROXYLATION.
 SQ SEQUENCE 11 AA; 1226 MW; 277AAC60B7232B58 CRC64;
 Query Match 23.5%; Score 24; DB 1; Length 11;
 Best Local Similarity 75.0%; Pred. No. 4.1e+02;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 6 KICH 9
 Db 6 KLCH 9
 RESULT 3
 CXL4_CONMR
 ID CXL4_CONMR STANDARD; PRT; 13 AA.
 AC P58810;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lambda/chi-conotoxin M1B (Chi-M1B).
 OS Conus marmoreus (Marble cone).
 OC Eukaryote; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=42752;
 RN [1]
 RP SEQUENCE, SYNTHESIS, AND STRUCTURE BY NMR.
 RC TISSUE-Venom;
 RA MEDLINE=21419681; PubMed=11528421;
 RA Sharpe I.A., Gehrmann J., Loughnan M.L., Adams D.A.,
 RA Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.,
 RA Lewis R.J.;
 RT "Two new classes of conopeptides inhibit the alpha1-adrenoceptor and
 RT noradrenaline transporter";
 RL Nat. Neurosci. 4:902-907(2001).
 CC -1- FUNCTION: Inhibits the neuronal noradrenaline transporter.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -1- PTM: Exists in two forms, due to cis-trans isomerization at His-
 CC 11-HYP-12.
 CC -1- MASS SPECTROMETRY: MW=1393.52; METHOD=Electrospray.
 CC -1- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
 DR PDB; ILEO; 03-APR-02.
 KW Neurotoxin; Toxin; Hydroxylation; 3D-structure.
 FT DISULFID 4 13
 FT DISULFID 5 10
 FT MOD_RES 12 12 HYDROXYLATION.
 SQ SEQUENCE 13 AA; 1382 MW; 277AAC376EAD2B58 CRC64;
 Query Match 23.5%; Score 24; DB 1; Length 13;
 Best Local Similarity 75.0%; Pred. No. 4.9e+02;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 6 KICH 9
 Db 8 KLCH 11

RESULT 4
 MLB_SQUAC
 ID MLB_SQUAC STANDARD; PRT; 16 AA.
 AC P01207;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Melanotropin beta.
 OS Squalus acanthias (Spiny dogfish).
 OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
 OX NCBI_TaxID=7797;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=75127390; PubMed=4375978;
 RA Bennett H.P.J., Lowry P.J., McMartin C., Scott A.P.;
 RT "Structural studies of alpha-melanocyte-stimulating hormone and a
 RT novel beta-melanocyte-stimulating hormone from the neurointermediate
 RT lobe of the pituitary of the dogfish Squalus acanthias";
 RL Blochem. J. 141:439-444(1974).
 CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
 CC -1- PIR; A01471; MTDIFS.
 KW Hormone.
 SQ SEQUENCE 16 AA; 1939 MW; 993AF43C8A640A0E CRC64;
 Query Match 23.5%; Score 24; DB 1; Length 16;
 Best Local Similarity 57.1%; Pred. No. 6e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 12 EYFGDF 18
 Db 4 DYKGFH 10
 RESULT 5
 MLB_SCYCA
 ID MLB_SCYCA STANDARD; PRT; 18 AA.
 AC P01206;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Melanotropin beta (Beta-MSH).
 OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
 OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
 OC Scyliorhinidae; Scyliorhinus.
 OX NCBI_TaxID=7830;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=75113445; PubMed=4452470;
 RA Love R.M., Pickering B.T.;
 RT "A beta-MSH in the pituitary gland of the spotted dogfish
 RT (Scyliorhinus canicula): Isolation and structure";
 RL Gen. Comp. Endocrinol. 24:398-404(1974).
 CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
 CC -1- PIR; A01470; MTDIFC.
 KW Hormone.
 SQ SEQUENCE 18 AA; 2195 MW; 353C93A30C314DF4 CRC64;
 Query Match 23.5%; Score 24; DB 1; Length 18;
 Best Local Similarity 44.4%; Pred. No. 6.7e+02;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 10 QIEYFGDF 18
 Db 2 ZIBYMGHF 10
 RESULT 6
 CXOB_CONMA
 ID CXOB_CONMA STANDARD; PRT; 25 AA.
 AC P05485;

DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Omega-conotoxin MV1b (SNX-159).
 OS Conus magus (Magus cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=6492;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=87299637; PubMed=2441741;
 RA Olivera B.M., Cruz L.J., de Santos V., Lechmanant G.W., Griffin D.,
 RA Zeikus R.D., McIntosh J.M., Galyean R., Varga J., Gray W.R.,
 RA Rivier J.E.;
 RT "Neuronal calcium channel antagonists. Discrimination between calcium
 channel subtypes using omega-conotoxin from Conus magus venom.";
 RL Biochemistry 26:2086-2090(1987).
 CC -1- FUNCTION: Omega-conotoxins act at presynaptic membranes, they bind
 CC and block voltage-sensitive calcium channels (VSCC).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -1- SIMILARITY: BELONGS TO THE O-SUPERFAMILY OF CONOTOXINS. OMEGA-TYPE
 CC FAMILY.
 DR PIR: B34115; B34115.
 DR PIR: JH0701; JH0701.
 DR HSP: P05484; IMVI.
 KW Presynaptic neurotoxin; Neurotoxin; Toxin; Calcium channel inhibitor;
 KW Anidation.
 FT DISULFID 1 16
 FT DISULFID 8 20
 FT DISULFID 15 25
 FT MOD_RES 25 25
 SQ SEQUENCE 25 AA; 2626 MW; E4B9CE5EFAA3734D CRC64;
 Query Match 23.58; Score 24; DB 1; Length 25;
 Best Local Similarity 50.0%; Pred. No. 9.4e+02;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 8 CHQIEY 13
 DB 8 CHRTSY 13

RESULT 7
 BTX_ATRBI STANDARD; PRT; 21 AA.
 AC P80163;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Bibtotoxin (BTX).
 OS Atractaspis bibroni (Southern bibron's) (Stiletto snake).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Atractaspididae; Atractaspis.
 OX NCBI_TaxID=8601;
 RN [1]
 RP SEQUENCE.
 RX TISSUE-Venom;
 RA Beckler A., Dowdle E.B., Hechler U., Kauser K., Donner P.,
 RA Schleuning W.-D.;
 RT "Bibtotoxin, a novel member of the endothelin/sarafotoxin peptide
 family, from the venom of the burrowing asp Atractaspis bibroni.";
 RL FEBS Lett. 315:100-103(1993).
 CC -1- FUNCTION: VASOCONSTRICTOR ACTIVITY. THESE TOXINS CAUSE CARDIAC
 CC ARREST PROBABLY AS A RESULT OF CORONARY VASOSPASM.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE ENDOTHELIN/SARAFOTOXIN FAMILY.
 DR PIR: S27039; S27039.
 DR HSP: P13208; 1SRB.
 DR InterPro: IPR001928; Endothln_tox.
 DR InterPro: IPR003642; Sara/bib_toxin.

DR Pfam: PF00322; endothelin; 1.
 DR PRINTS; PR00365; ENDOTHELIN.
 DR PRODOM; PD004740; Sara/bib_toxin; 1.
 DR SMART; SM00272; END; 1.
 DR PROSITE; PS00270; ENDOTHELIN; 1.
 KW Vasoconstrictor; Toxin.
 FT DISULFID 1 15
 FT DISULFID 3 11
 SQ SEQUENCE 21 AA; 2511 MW; 83A5DFB81D036AE2 CRC64;
 Query Match 21.6%; Score 22; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 CHQ 10
 DB 15 CHQ 17

RESULT 8
 SRTD_ATREN STANDARD; PRT; 21 AA.
 AC P13211;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sarafotoxin-D (S6D) (SRTX-D).
 OS Atractaspis enadensis (Israeli burrowing asp).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Atractaspididae; Atractaspis.
 OX NCBI_TaxID=8600;
 RN [1]
 RP SEQUENCE.
 RX TISSUE-Venom;
 RX MEDLINE=90033283; PubMed=2509240;
 RA Bclah A., Wellberg Z., Fleminger G.;
 RT "SRTX-d, a new native peptide of the endothelin/sarafotoxin family.";
 RL FEBS Lett. 256:1-3(1989).
 CC -1- FUNCTION: VASOCONSTRICTOR ACTIVITY. THESE TOXINS CAUSE CARDIAC
 CC ARREST PROBABLY AS A RESULT OF CORONARY VASOSPASM.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: LD(50) IS 0.35 MG/KG BY INTRAVENOUS INJECTION.
 CC -1- SIMILARITY: BELONGS TO THE ENDOTHELIN/SARAFOTOXIN FAMILY.
 DR PIR: S06145; S06145.
 DR HSP: P13208; 1SRB.
 DR InterPro: IPR001928; Endothln_tox.
 DR InterPro: IPR003642; Sara/bib_toxin.
 DR Pfam: PF00322; endothelin; 1.
 DR PRINTS; PR00365; ENDOTHELIN.
 DR PRODOM; PD004740; Sara/bib_toxin; 1.
 DR SMART; SM00272; END; 1.
 DR PROSITE; PS00270; ENDOTHELIN; 1.
 KW Vasoconstrictor; Toxin.
 FT DISULFID 1 15
 FT DISULFID 3 11
 SQ SEQUENCE 21 AA; 2596 MW; 83A8A04D1D536AE2 CRC64;
 Query Match 21.6%; Score 22; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 CHQ 10
 DB 15 CHQ 17

RESULT 9
 ANDT_ANDAU STANDARD; PRT; 25 AA.
 ID ANDT_ANDAU
 AC P56684; P81616;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Androctonin.
OS Androctonus australis hector (Sahara scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Butioidae; Butiidae; Androctonus.
OX NCBI_TaxID=70175;
RN [1]
RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
RC TISSUE-Hemolymph;
RX MEDLINE=97094646; PubMed=8939880;
RA Elret-Sabatier L., Loew D., Goyffon M., Fehlbaum P., Hoffmann J.A.,
van Dorsselaer A., Bulet P.;
RT "Characterization of novel cysteine-rich antimicrobial peptides from
scorpion blood.";
RL J. Biol. Chem. 271:29537-29544(1996).
RN [2]
RP SYNTHESIS OF D-AMINO ACID ENANTOMER, AND CHARACTERIZATION.
RX MEDLINE=20115101; PubMed=10642525;
RA Hetru C., Letellier L., Oren Z., Hoffmann J.A., Shai Y.;
RT "Androctonin, a hydrophilic disulphide-bridged non-haemolytic
anti-microbial peptide: a plausible mode of action.";
RL Biochem. J. 345:653-664(2000).
RN [3]
RP STRUCTURE BY NMR.
RX MEDLINE=20025109; PubMed=10563585;
RA Mandard N., Sy D., Maufrais C., Bonmatin J.M., Bulet P., Hetru C.,
Vovelle F.;
RT "Androctonin, a novel antimicrobial peptide from scorpion Androctonus
australis: solution structure and molecular dynamics simulations in
the presence of a lipid monolayer.";
RL J. Biomol. Struct. Dyn. 17:367-380(1999).
CC -1- FUNCTION: ACTIVE AGAINST BOTH BACTERIA (GRAM-POSITIVE AND GRAM-
NEGATIVE) AND FILAMENTOUS FUNGI. ACTS ON THE MEMBRANE OF THE
BACTERIAL CELLS. IT DESTABILIZE A MEMBRANE BY MODIFYING ITS
PROPERTIES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC PDB: 1C26; 12-JAN-00.
DR Antibiotic; Fungicide; 3D-structure.
FT DISULFID 4 20
FT SEQUENCE 25 AA; 3081 MW; D4183D6EA5AA50AA CRC64;
SQ
Query Match 21.6%; Score 22; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 7 ICHQIE 12
DB 3 VCRQIK 8
RESULT 10
GON2_CHEPR STANDARD; PRT; 10 AA.
AC P80678;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gonadolibirin II (Gonadotropin-releasing hormone II) (GNRH-II)
DE (Luliberin II).
OS Chelyosoma productum.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Corellidae; Chelyosoma.
OX NCBI_TaxID=71177;
RN [1]
RP SEQUENCE.
RX MEDLINE=96413669; PubMed=8816823;
RA Powell J.F.F., Reska-Skinner S.M., Prakash M.O., Fischer W.H.,
Park M., Rivier J.E., Craig A.G., Mackie G.O., Sherwood N.M.;
RT "Two new forms of gonadotropin-releasing hormone in a protochordate
and the evolutionary implications.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:10461-10464(1996).

CC -1- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
FOLLICLE-STIMULATING HORMONES.
CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: GNRH NEURONS LIE WITHIN BLOOD SINUSES CLOSE TO
THE GONODUCTS AND GONADS IN BOTH JUVENILES AND ADULTS, IMPLYING
THAT THE NEUROPEPTIDE IS RELEASED INTO THE BLOODSTREAM.
CC -1- MASS SPECTROMETRY: MW=1117.52; METHOD=MALDI.
CC -1- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR InterPro: IPR002012; Gnrh.
KW Hormone; Amidation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 6 6 INTERCHAIN.
FT MOD_RES 10 10 AMIDATION (BY SIMILARITY).
SQ SEQUENCE 10 AA; 1135 MW; 284B38D1EEB735A3 CRC64;
Query Match 20.6%; Score 21; DB 1; Length 10;
Best Local Similarity 42.9%; Pred. No. 1.1e+03;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 3 QEAKICH 9
DB 1 QHWSLCH 7
RESULT 11
TIN2_HOPTI STANDARD; PRT; 12 AA.
AC P82652;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tigerin-2.
OS Hoplobatrachus tigerinus (Indian bull frog) (Rana tigerina).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae;
OC Hoplobatrachus.
OX NCBI_TaxID=103373;
RN [1]
RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND DISULFIDE BONDS.
RC TISSUE-Skin.
RX PubMed=11031261;
RA Purna Sai K., Jagannadham M.V., Vairamani M., Raju N.P.,
Devi A.S., Nagaraj R., Sitarum N.;
RT "Tigerinins: novel antimicrobial peptides from the Indian frog Rana
tigerina.";
RL J. Biol. Chem. 276:2701-2707(2001).
CC -1- FUNCTION: ANTIBACTERIAL ACTIVITY AGAINST B. SUBTILIS, E. COLI,
S. AUREUS, M. LUTEUS, P. PUTIDA AND S. CEREVISIAE.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: SKIN.
CC -1- MASS SPECTROMETRY: MW=1368; METHOD=MALDI.
KW Amphibian skin; Antibiotic; Amidation.
FT DISULFID 3 11
FT MOD_RES 12 12 AMIDATION.
SQ SEQUENCE 12 AA; 1369 MW; C2627D8EFD37605D CRC64;
Query Match 20.6%; Score 21; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 ICH 9
DB 10 ICH 12
RESULT 12
TPIS_PINPS STANDARD; PRT; 17 AA.
ID TPIS_PINPS
AC P81666;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update)
Triosephosphate isomerase, cytosolic (EC 5.3.1.1) (TIM) (Fragments).
Pinus pinaster (Maritime pine).
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
NCBI_TaxID=71647;
[1]
SEQUENCE.
PC TISSUE=Needle;
RX MEDLINE=9274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Prigiero J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
proteins";
RL Electrophoresis 20:1098-1108(1999).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate -> glyceralone
phosphate.
CC -1- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- INDUCTION: BY WATER STRESS.
CC -1- MISCELLANEOUS: IN PLANTS, THERE ARE TWO TYPES OF TPIS, CYTOSOLIC
AND PLASTID.
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN
(SPOT N139) IS: 5.9, ITS MW IS: 24 kDa.
CC -1- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
DR InterPro: IPR000652; Triophos-ismrse.
DR PROSITE: PS00171; TIM; PARTIAL.
KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
Pentose shunt.
FT NON_TER 1 1
FT NON_CONS 9 10
FT NON_TER 17 17
SQ SEQUENCE 17 AA; 2053 MW; 183DB41757AF13CB CRC64;

Query Match 20.6%; Score 21; DB 1; Length 17;
Best Local Similarity 20.0%; Pred. No. 1.9e+03;
Matches 2; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 7 ICHQIEYFG 16
: : : : :
Db 4 VCYEQLFFVG 13

RESULT 13
REV_HV2D2 STANDARD; PRT; 21 AA.
AC PL5830;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE REV protein (Anti-repression transactivator protein) (ART/TRS)
DE (fragment).
GN REV.
OS Human immunodeficiency virus type 2 (isolate D205,7) (HIV-2).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11716;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90081881; PubMed=2594088;
RA Dietrich U., Adamski M., Kreutz R., Seipp A., Kuehnel H.,
RA Ruebsamen-Waigmann H.;
RT "A highly divergent HIV-2-related isolate.";
RL Nature 342:948-950(1989).
CC -1- FUNCTION: REV APPEARS TO ACT POST-TRANSCRIPTIONALLY TO RELIEVE
NEGATIVE REPRESSION OF GAG AND ENV PRODUCTION.
CC -1- SUBCELLULAR LOCATION: Nuclear; accumulates in the nucleoli.
CC -1- PTM: PHOSPHOPROTEIN WHOSE STATE OF PHOSPHORYLATION IS MEDIATED
BY A SPECIFIC SERINE KINASE ACTIVITY PRESENT IN THE NUCLEUS.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch.

DR EMBL: X61240; NOT_ANNOTATED_CDS.
DR FIR: S08441; S08441.
DR HIV: X16109; REV\$2D205.
DR InterPro: IPR000625; REV_protein.
DR Pfam: PF00424; REV; 1.
KW Transcription regulation; AIDS; Phosphorylation; Nuclear protein.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2503 MW; E620E225CC5BFF24 CRC64;

Query Match 20.6%; Score 21; DB 1; Length 21;
Best Local Similarity 37.5%; Pred. No. 2.4e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 QOEAKICH 9
: : : : :
Db 9 QKELRLIH 16

RESULT 14
DP20_CAEBR STANDARD; PRT; 22 AA.
ID DP20_CAEBR
AC P51558;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein dpy-20 (Fragment).
GN dpy-20.
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95287857; PubMed=7770042;
RA Clark D.V., Suleman D.S., Beckenbach K.A., Gilchrist E.J.,
RA Baillie D.L.;
RT "Molecular cloning and characterization of the dpy-20 gene of
Caenorhabditis elegans.";
RL Mol. Gen. Genet. 247:367-378(1995).
CC -1- FUNCTION: INVOLVED IN CUTICLE FUNCTION AND IS ESSENTIAL FOR
NORMAL MORPHOLOGICAL DEVELOPMENT (BY SIMILARITY).
CC TRANSFAC: T04319; -
KW Cuticle.
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2529 MW; A11FB717AD367F69 CRC64;

Query Match 20.6%; Score 21; DB 1; Length 22;
Best Local Similarity 42.9%; Pred. No. 2.5e+03;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOEAKI 7
: : : : :
Db 15 QQQQSQL 21

RESULT 15
FUC1_RAT STANDARD; PRT; 22 AA.
ID FUC1_RAT
AC P80347;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last annotation update)
DE Fucinin 1 (Fucosyltransferase inhibitor 1) (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

RP SEQUENCE.
 RC STRAIN-Sprague-Dawley; TISSUE-Small intestine mucosa;
 RA MEDLINE=94357231; PubMed=8076650;
 RA Ruggiero-Lopez D., Manioc C., Geourjon C., Louisot P., Martin A.;
 RT "Purification and partial amino acid sequence of fucinin, an
 RT endogenous inhibitor of fucosyltransferase activities.";
 RL Eur. J. Biochem. 224:47-55(1994).
 CC -1- FUNCTION: HAS A ROLE IN THE PHYSIOLOGICAL REGULATION OF
 CC FUCOSYLATION PROCESSES.
 CC -1- SUBUNIT: OLIGOMER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: SOME, WITH HUMAN SET/PHAPII PROTEIN.
 FT NON_TER 22
 SQ SEQUENCE 22 AA; 2393 MW; 0A12574A68A8E8A9 CRC64;
 Query Match 20.6%; Score 21; DB 1; Length 22;
 Best Local Similarity 80.0%; Pred. No. 2.5e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QQEA 5
 Db 12 EQEA 16

Search completed: April 23, 2003, 13:43:49
 Job time : 4.75169 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:36:43 ; Search time 9.40449 Seconds
(without alignments)
183.999 Million cell updates/sec

Title: US-09-836-073-4
Perfect score: 102
Sequence: 1 QQQEAKICHQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 4984

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	30.4	22	PH1359	Ig heavy chain DJ
2	30	29.4	25	A60286	heat-stable serine
3	29	28.4	18	PH1368	Ig heavy chain DJ
4	28	27.5	21	S78574	protein kinase C 1
5	27	26.5	12	PH0771	T-cell receptor be
6	27	26.5	14	PH1626	Ig H chain V-D-J r
7	27	26.5	15	PH0789	T-cell receptor al
8	27	26.5	19	A05305	hemoglobin beta-2
9	27	26.5	21	S78575	protein kinase C 1
10	27	26.5	22	PH1325	Ig heavy chain DJ
11	27	26.5	23	PH1725	Ig heavy chain V r
12	27	26.5	24	PH1696	Ig heavy chain V r
13	27	26.5	25	C57001	endo-1,4-beta-xyla
14	26	25.5	14	PH1598	Ig H chain V-D-J r
15	26	25.5	16	E53284	T-cell receptor be
16	26	25.5	18	PH1629	Ig H chain V-D-J r
17	26	25.5	19	B53145	high conductance c
18	26	25.5	22	I77373	gene N-ras protein
19	26	25.5	24	B53524	ubiquinol-cytochro
20	26	25.5	25	PH1700	Ig heavy chain V r
21	25	24.5	12	G64003	hypothetical prote
22	25	24.5	15	PH1366	Ig heavy chain DJ
23	25	24.5	17	A61211	anantin - Streptom
24	25	24.5	24	PT0258	Ig heavy chain CDR
25	25	24.5	25	PH1686	Ig heavy chain V r
26	24	23.5	12	S57570	T cell receptor V-
27	24	23.5	14	S57569	T cell receptor V-
28	24	23.5	14	S57638	T cell receptor V-
29	24	23.5	16	MTDFBS	melanotropin beta

30	24	23.5	18	1	MTDFBC	melanotropin beta
31	24	23.5	21	2	PH1730	Ig heavy chain V r
32	24	23.5	21	2	T07683	proteinase inhibit
33	24	23.5	22	2	PH1678	Ig heavy chain V r
34	24	23.5	22	2	PH1679	Ig heavy chain V r
35	24	23.5	23	2	PH1681	Ig heavy chain V r
36	24	23.5	23	2	PH1682	Ig heavy chain V r
37	24	23.5	23	2	PH1694	Ig heavy chain V r
38	24	23.5	23	2	PH1707	Ig heavy chain V r
39	24	23.5	23	2	PH1722	Ig heavy chain V r
40	24	23.5	23	2	PH1724	Ig heavy chain V r
41	24	23.5	23	2	PH1727	Ig heavy chain V r
42	24	23.5	23	2	PH1723	Ig heavy chain V r
43	24	23.5	24	2	PH1683	Ig heavy chain V r
44	24	23.5	24	2	PH1685	Ig heavy chain V r
45	24	23.5	24	2	PH1698	Ig heavy chain V r

ALIGNMENTS

RESULT 1

PH1359
Ig heavy chain DJ region (clone C178-121) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PH1359
R;Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Roveta, G.
J. Exp. Med. 176, 1577-1581, 1992
A;Title: Predominance of fetal type DJH joining in young children with B precursor :
A;Reference number: PH1302; MUID:93094761; PMID:1460419
A;Accession: PH1359
A;Molecule type: DNA
A;Residues: 1-22 <WAS>
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotrimer; immunoglobulin

Query Match 30.4%; Score 31; DB 2; Length 22;
Best Local Similarity 44.4%; Pred. No. 1.5e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 8 CHQIEYFG 16
|::|||:
Db 6 CYENYYYG 14

RESULT 2

A60286
heat-stable serine proteinase (EC 3.4.21.-) - Thermomonospora fusca (strain YX) (fr:
N;Alternate names: YX-proteinase
C;Species: Thermomonospora fusca
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Dec-1996
C;Accession: A60286
R;Kristjansson, M.M.; Kinsella, J.E.
Int. J. Pept. Protein Res. 36, 201-207, 1990
A;Title: Heat stable proteinase from Thermomonospora fusca. Characterization as a s:
A;Reference number: A60286; MUID:91107200; PMID:2132918
A;Accession: A60286
A;Molecule type: protein
A;Residues: 1-25 <KRI>
C;Keywords: extracellular protein; hydrolase; serine proteinase

Query Match 29.4%; Score 30; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 13 YFGDF 18
||||:
Db 10 YFGNY 15

RESULT 3
PH1368

Ig heavy chain DJ region (clone C111-112) - human (fragment)

C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C;Accession: PH1368
R;Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.

J. Exp. Med. 176, 1577-1581, 1992
A;Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma

A;Reference number: PH1302; MUID:93094761; PMID:1460419
A;Accession: PH1368

A;Molecule type: DNA
A;Residues: 1-18 <WAS>

C;Superfamily: Immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotrimer; immunoglobulin

Query Match 28.4%; Score 29; DB 2; Length 18;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 IEYVFG 16
: : : : :
: : : : :

Db 5 MEYVYG 10
: : : : :
: : : : :

RESULT 4
S78574

protein kinase C inhibitor KCIP-1 isoform gamma - sheep (fragment)

C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 21-Nov-1998

C;Accession: S78574
R;Martin, H.; Patel, Y.; Jones, D.; Howell, S.; Robinson, K.; Attkin, A.

FEBS Lett. 331, 296-303, 1993
A;Title: Antibodies against the major brain isoforms of 14-3-3 protein. An antibody specific for the 14-3-3 protein

A;Reference number: S38299; MUID:93387487; PMID:8375512
A;Accession: S78574

A;Molecule type: protein
A;Residues: 1-21 <MAK>

A;Note: The acetylated initiator methionine is removed and the Val undergoes a further modification

C;Superfamily: 14-3-3 protein
C;Keywords: acetylated amino end

F;1/Modified site: acetylated amino end (Val) #status experimental

Query Match 27.5%; Score 28; DB 2; Length 21;
Best Local Similarity 41.1%; Pred. No. 4.4e+02;

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 QEAKICHQIEVY 14
: : : : :
: : : : :

Db 8 QKARLAQERY 19
: : : : :
: : : : :

RESULT 5
PH0771

T-cell receptor beta chain (PE5.1.1) - mouse (fragment)

C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999

C;Accession: PH0771
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.

J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex

allelic exclusion and antigen-specific repertoire.

A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0771

A;Molecule type: mRNA
A;Residues: 1-12 <CAS>

A;Cross-references: EMBL:X60865; NID:953624; PIDN:CA43255.1; PID:953625

A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match 26.5%; Score 27; DB 2; Length 12;
Best Local Similarity 71.4%; Pred. No. 3.7e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYVFG 16
: : : : :
: : : : :

Db 6 QYEQYFG 12
: : : : :
: : : : :

RESULT 6
PH1626

Ig H chain V-D-J region (clone B-less 118) - mouse (fragment)

C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999

C;Accession: PH1626
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.

J. Exp. Med. 178, 317-329, 1993
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less

A;Reference number: PH1580; MUID:93301609; PMID:8315387
A;Accession: PH1626

A;Molecule type: DNA
A;Residues: 1-14 <LEV>

A;Experimental source: bone marrow pre-B lymphocyte
C;Keywords: immunoglobulin

Query Match 26.5%; Score 27; DB 2; Length 14;
Best Local Similarity 40.0%; Pred. No. 4.3e+02;

Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 CHQIEYVFG 17
: : : : :
: : : : :

Db 1 CMRAXXYGD 10
: : : : :
: : : : :

RESULT 7
PH0789

T-cell receptor alpha chain (E22 V-alpha-4.delta-7R) - mouse (fragment)

C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C;Accession: PH0789
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.

J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex

allelic exclusion and antigen-specific repertoire.

A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0789

A;Molecule type: mRNA
A;Residues: 1-15 <CAS>

A;Cross-references: EMBL:X60894

A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match 26.5%; Score 27; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 4.6e+02;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 HQIEYVFG 16
: : : : :
: : : : :

Db 8 HGLQTYFG 15
: : : : :
: : : : :

RESULT 8
A05305

hemoglobin beta-2 chain - Indian spiny-tailed lizard (fragment)

C;Species: Uromastix hardwickii (Indian spiny-tailed lizard)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 01-Aug-1997

C;Accession: A05305
R;Naqvi, S.; Zaidi, Z.H.; von Bahr-Lindstrom, H.; Carlquist, M.; Jornvall, H.

FEBS Lett. 162, 290-295, 1983
A;Reference number: A91314; MUID:84029159; PMID:6628672

A;Accession: A05305
A;Molecule type: protein

A;Residues: 1-19 <NAQ>
C;Superfamily: globin; globin homology

C;Keywords: erythrocyte; oxygen carrier

Query Match 26.5%; Score 27; DB 2; Length 19;
Best Local Similarity 80.0%; Pred. No. 5.8e+02;

QY 10 QIEYVFG 16
: : : : :
: : : : :

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 14 YFGDF 18
:||||
Db 1 FFGDF 5

RESULT 9

S78575
protein kinase C inhibitor KCIP-1 isoform eta - sheep (fragment)
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 21-Nov-1998
C:Accession: S78575
R:Martin, H.; Patel, Y.; Jones, D.; Howell, S.; Robinson, K.; Altken, A.
FEBS Lett. 331, 296-303, 1993
A:Title: Antibodies against the major brain isoforms of 14-3-3 protein. An antibody spec
A:Reference number: S38299; MUID:93387487; PMID:8375512
A:Accession: S78575
A:Molecule type: protein
A:Residues: 1-21 <MAX>
A:Note: the acetylated initiator methionine is removed and the Gly undergoes a further a
C:Superfamily: 14-3-3 protein
C:Keywords: acetylated amino end
F:1/Modified site: acetylated amino end (Gly) #status experimental

Query Match 26.5%; Score 27; DB 2; Length 21;
Best Local Similarity 41.7%; Pred. No. 6.4e+02;

Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 QKAKICHOIEY 14
|:::|
Db 8 QRARLAQAERY 19

RESULT 10

PH1325
Ig heavy chain DJ region (clone C199-121) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PH1325
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymph
A:Reference number: PH1302; MUID:93094761; PMID:1460419
A:Accession: PH1325
A:Molecule type: DNA
A:Residues: 1-22 <WAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 26.5%; Score 27; DB 2; Length 22;
Best Local Similarity 66.7%; Pred. No. 6.7e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 11 IEYFYG 16
|:::|
Db 9 IHYYYG 14

RESULT 11

PH1725
Ig heavy chain V region (clone GCC-5) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385
A:Accession: PH1725
A:Molecule type: mRNA
A:Residues: 1-23 <MCH>
A:Experimental source: B cell

A:Note: the authors translated the codon ACA for residue 13 as Ala
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 26.5%; Score 27; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 7e+02;

Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 10 QIEYFQDF 18
:||||
Db 12 ETRYVGSY 20

RESULT 12

PH1696
Ig heavy chain V region (clone NP-7-9) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1696
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385
A:Accession: PH1696
A:Molecule type: mRNA
A:Residues: 1-24 <MCH>
A:Experimental source: B cell
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 26.5%; Score 27; DB 2; Length 24;
Best Local Similarity 42.9%; Pred. No. 7.3e+02;

Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 10 QIEYFYG 16
:||||
Db 12 EVAYYIG 18

RESULT 13

C57001
endo-1,4-beta-xylanase (EC 3.2.1.8) 4 - Streptomyces roseiscleroticus (fragment)
C:Species: Streptomyces roseiscleroticus
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 23-May-1997
C:Accession: C57001
R:Grabski, A.C.; Forrester, I.T.; Patel, R.; Jeffries, T.W.
Protein Expr. Purif. 4, 120-129, 1993
A:Title: Characterization and N-terminal amino acid sequences of beta-(1-4)endoxyla
A:Reference number: A57001; MUID:93229899; PMID:8471845
A:Accession: C57001
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-25 <GRA>
A:Experimental source: strain NRRL B-11019
A:Note: sequence extracted from NCBI backbone (NCBIP:130009)
C:Function:
A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
A:Pathway: xylan degradation
C:Superfamily: Streptomyces endo-1,4-beta-xylanase A; Streptomyces endo-1,4-beta-xy
C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradati

Query Match 26.5%; Score 27; DB 2; Length 25;
Best Local Similarity 71.4%; Pred. No. 7.7e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 10 QIEYFYG 16
|:::|
Db 11 QSGYFYG 17

RESULT 14

PH1598
Ig H chain V-D-J region (wild-type clone 306) - mouse (fragment)

C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C:Accession: PH1598
R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A:Reference number: PH1580; MUID:93301609; PMID:8315387
A:Accession: PH1598
A:Molecule type: DNA
A:Residues: 1-14 <LEV>
A:Experimental source: bone marrow pre-B lymphocyte
C:Keywords: immunoglobulin

Query Match 25.5%; Score 26; DB 2; Length 14;
Best Local Similarity 50.0%; Pred. NO. 6.3e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 8 CHQIEYFGD 17
| : ||||
Db 1 CARDGYGGD 10

RESULT 15
E53284
T-cell receptor beta 2 chain J region, Jbeta2.2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: E53284
R:Harindranath, N.; Alexander, C.B.; Mage, R.G.
Mol. Immunol. 28, 881-888, 1991
A:Title: Evolutionarily conserved organization and sequences of germline diversity and
A:Reference number: A53284; MUID:91342695; PMID:1678859
A:Accession: E53284
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-16 <HAR>
A:Cross-references: GB:S60737; NID:g233916; PIDN:AA19521.1; PID:g233921
A:Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBIP:60743)
C:Keywords: T-cell receptor

Query Match 25.5%; Score 26; DB 2; Length 16;
Best Local Similarity 66.7%; Pred. NO. 7.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 12 EYVYFGD 17
| : ||||
Db 4 QLYFGD 9

Search completed: April 23, 2003, 13:48:54
Job time : 10.4045 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:43:09 ; Search time 9.91011 Seconds
(without alignments)
145.542 Million cell updates/sec

Title: US-09-836-073-4

Perfect score: 102

Sequence: 1 QQEAKICHQIEYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 61792

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap:*
- 2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pap:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap:*
- 7: /cgn2_6/ptodata/2/pubpaa/PTCTUS_PUBCOMB.pap:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	100.0	18	US-09-836-073-4	Sequence 4, Appli
2	90	88.2	18	US-09-836-073-5	Sequence 5, Appli
3	87	85.3	17	US-09-836-073-13	Sequence 13, Appl
4	87	85.3	18	US-09-836-073-1	Sequence 1, Appli
5	87	85.3	18	US-09-836-073-14	Sequence 14, Appl
6	81	79.4	18	US-09-836-073-9	Sequence 9, Appli
7	79	77.5	18	US-09-836-073-11	Sequence 11, Appl
8	79	77.5	18	US-09-836-073-12	Sequence 12, Appl
9	78	76.5	18	US-09-836-073-10	Sequence 10, Appl
10	76.5	75.0	19	US-09-836-073-16	Sequence 16, Appl
11	75	73.5	18	US-09-836-073-2	Sequence 2, Appli
12	75	73.5	18	US-09-836-073-3	Sequence 3, Appli
13	72	70.6	18	US-09-836-073-7	Sequence 7, Appli
14	72	70.6	18	US-09-836-073-15	Sequence 15, Appl
15	71	69.6	18	US-09-836-073-8	Sequence 8, Appli
16	62	60.8	16	US-09-836-073-19	Sequence 19, Appl
17	56	54.9	18	US-09-836-073-6	Sequence 6, Appli
18	44	43.1	18	US-09-836-073-17	Sequence 17, Appl
19	38.5	37.7	18	US-09-836-073-18	Sequence 18, Appl

20	30.5	29.9	23	9	US-09-813-153-291	Sequence 291, App
21	30	29.4	25	9	US-10-097-065-405	Sequence 405, App
22	29	28.4	14	8	US-08-424-550B-455	Sequence 455, App
23	29	28.4	17	10	US-09-864-761-47213	Sequence 47213, A
24	29	28.4	19	9	US-10-153-159-6	Sequence 6, Appli
25	29	28.4	19	9	US-10-153-159-51	Sequence 51, Appl
26	29	28.4	19	9	US-10-153-159-53	Sequence 53, Appl
27	29	28.4	19	9	US-10-153-159-54	Sequence 54, Appl
28	29	28.4	19	9	US-10-153-176-6	Sequence 6, Appli
29	29	28.4	19	9	US-10-153-176-51	Sequence 51, Appl
30	29	28.4	19	9	US-10-153-176-53	Sequence 53, Appl
31	29	28.4	19	9	US-10-153-176-54	Sequence 54, Appl
32	29	28.4	20	9	US-09-986-480-269	Sequence 269, App
33	28	27.5	15	10	US-09-073-009-51	Sequence 51, Appl
34	28	27.5	15	10	US-09-023-588-51	Sequence 51, Appl
35	28	27.5	15	10	US-09-793-306-51	Sequence 51, Appl
36	28	27.5	18	9	US-10-084-813-139	Sequence 139, App
37	28	27.5	18	9	US-10-084-813-140	Sequence 140, App
38	28	27.5	18	9	US-10-084-813-141	Sequence 141, App
39	28	27.5	21	9	US-09-880-748-3010	Sequence 3010, Ap
40	28	27.5	25	10	US-09-864-761-43809	Sequence 43809, A
41	27	26.5	15	9	US-09-989-919-123	Sequence 123, App
42	27	26.5	17	10	US-09-864-761-33337	Sequence 33337, A
43	27	26.5	18	10	US-09-840-009-13	Sequence 13, Appl
44	27	26.5	19	9	US-10-153-159-49	Sequence 49, Appl
45	27	26.5	19	9	US-10-153-159-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836.073
; CURRENT FILING DATE: 2002-10-24
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match 100.0%; Score 102; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.5e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QQEAKICHQIEYFGDF 18

Db 1 QQEAKICHQIEYFGDF 18

RESULT 2

US-09-836-073-5
; Sequence 5, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836.073

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-5

Query Match 88.2%; Score 90; DB 9; Length 18;
Best Local Similarity 88.9%; Pred. No. 7.3e-08;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 QQQEAKICHOIEYFGDF 18
||||| |||||||
Db 1 QQQEAKICHOIEYFGDF 18

RESULT 3
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match 85.3%; Score 87; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
||||| |||||||
Db 3 EAKICHOIEYFGDF 17

RESULT 4
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match 85.3%; Score 87; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
||||| |||||||
Db 4 EAKICHOIEYFGDF 18

RESULT 5
US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14

Query Match 85.3%; Score 87; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
||||| |||||||
Db 4 EAKICHOIEYFGDF 18

RESULT 6
US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match 79.4%; Score 81; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17
||||| |||||||
Db 4 EAKICHOIEYFGD 17

RESULT 7
US-09-836-073-11


```
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11:
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match      77.5%; Score 79; DB 9; Length 18;
Best Local Similarity 93.3%; Pred. No. 3.9e-06;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 EAKICHQIEYFGDF 18
Db 4 EAKICHQIEYFGDF 18
|||||
|||||

RESULT 8
US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12:
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match      77.5%; Score 79; DB 9; Length 18;
Best Local Similarity 93.3%; Pred. No. 3.9e-06;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 EAKICHQIEYFGDF 18
Db 4 EAKICHQIEYFGDF 18
|||||
|||||
```

```
US-09-836-073-12

Query Match      77.5%; Score 79; DB 9; Length 18;
Best Local Similarity 93.3%; Pred. No. 3.9e-06;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 EAKICHQIEYFGDF 18
Db 4 EAKICHQIEYFGDF 18
|||||
|||||

RESULT 9
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
```

```
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10:
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match      76.5%; Score 78; DB 9; Length 18;
Best Local Similarity 93.3%; Pred. No. 5.6e-06;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 EAKICHQIEYFGDF 18
Db 4 EAKICHQIEYFGDF 18
|||||
|||||

RESULT 10
US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16:
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match      75.0%; Score 76.5; DB 9; Length 19;
Best Local Similarity 93.8%; Pred. No. 1e-05;
Matches 15; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
```

```
Qy 4 EAKICHQI-EYFGDF 18
Db 4 EAKICHQIEYFGDF 19
|||||
|||||

RESULT 11
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2:
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match      73.5%; Score 75; DB 9; Length 18;
Best Local Similarity 86.7%; Pred. No. 1.7e-05;
```

Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
|||||

Db 4 EAKICHOIEYFGDF 18
|||||

RESULT 12

US-09-836-073-3

; Sequence 3, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; PRIOR FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-3

Query Match

Best Local Similarity 73.5%; Score 75; DB 9; Length 18;

Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
:|||||:|||||

Db 4 EAKICHOIEYFGDF 18
:|||||:|||||

RESULT 13

US-09-836-073-7

; Sequence 7, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; PRIOR FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-7

Query Match

Best Local Similarity 70.6%; Score 72; DB 9; Length 18;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGD 17
|||||

Db 4 EAKICHOIEYQGD 17
|||||

RESULT 14

US-09-836-073-15

; Sequence 15, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Xenopus
US-09-836-073-15

Query Match

Best Local Similarity 70.6%; Score 72; DB 9; Length 18;

Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
:|||||

Db 4 DTRICEQIEYFGDF 18
:|||||

RESULT 15

US-09-836-073-8

; Sequence 8, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-8

Query Match

Best Local Similarity 69.6%; Score 71; DB 9; Length 18;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
|||||

Db 4 EAKICHOIEQOFGDF 18
|||||

Search completed: April 23, 2003, 13:52:09

Job time : 10.9101 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:38:28 ; Search time 8.49438 Seconds
(without alignments)
62.349 Million cell updates/sec

Title: US-09-836-073-4
Perfect score: 102
Sequence: 1 QOQEKICHQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 147762

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA.*
1: /cgn2_5/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	85.3	18	4	US-09-316-630-3
2	87	85.3	18	4	US-09-316-630-4
3	35	34.3	24	4	US-09-082-279B-1195
4	35	34.3	24	4	US-09-315-304B-1195
5	30	29.4	7	1	US-08-166-930-15
6	30	29.4	7	2	US-08-727-045A-15
7	30	29.4	7	2	US-09-408-172-15
8	30	29.4	11	2	US-08-618-696-20
9	30	29.4	11	2	US-08-211-312-9
10	30	29.4	11	3	US-09-033-753-20
11	30	29.4	11	3	US-08-472-285-9
12	30	29.4	11	4	US-08-472-929-9
13	29	28.4	10	2	US-08-618-696-11
14	29	28.4	10	3	US-09-033-753-11
15	29	28.4	11	2	US-08-618-696-7
16	29	28.4	11	3	US-09-033-753-7
17	29	28.4	14	2	US-08-637-759B-113
18	29	28.4	14	3	US-08-871-355A-113
19	29	28.4	14	4	US-09-201-945-113
20	29	28.4	14	4	US-09-298-924-55
21	29	28.4	14	4	US-08-469-260A-455
22	29	28.4	15	4	US-08-743-168B-29
23	29	28.4	15	5	PCT-US96-10435-29
24	29	28.4	25	1	US-08-238-163-6
25	28	27.5	10	2	US-08-618-696-2
26	28	27.5	10	2	US-08-618-696-10
27	28	27.5	10	3	US-09-033-753-2

28	27.5	10	3	US-09-033-753-10	Sequence 10, Appl
29	27.5	11	2	US-08-618-696-1	Sequence 1, Appl
30	27.5	11	2	US-08-618-696-6	Sequence 6, Appl
31	27.5	11	2	US-08-618-696-16	Sequence 16, Appl
32	27.5	11	2	US-08-618-696-19	Sequence 19, Appl
33	27.5	11	3	US-09-033-753-1	Sequence 1, Appl
34	27.5	11	3	US-09-033-753-6	Sequence 6, Appl
35	27.5	11	3	US-09-033-753-16	Sequence 16, Appl
36	27.5	11	3	US-09-033-753-19	Sequence 19, Appl
37	27.5	19	4	US-09-441-502B-61	Sequence 61, Appl
38	27.5	20	4	US-08-505-250-8	Sequence 8, Appl
39	27.5	20	4	US-08-505-250-8	Sequence 8, Appl
40	27.5	21	2	US-08-825-349-1	Sequence 1, Appl
41	27.5	22	6	5281520-43	Patent No. 5281520
42	27.5	24	4	US-09-082-279B-1199	Sequence 1199, Ap
43	27.5	24	4	US-07-963-329A-73	Sequence 73, Appl
44	27.5	24	4	US-09-315-304B-1199	Sequence 1199, Ap
45	27.5	24	5	PCT-US92-09443A-73	Sequence 73, Appl

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316.630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817.953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321.427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086.527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3
Query Match 85.3%; Score 87; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 EAKICHOIEYFGDF 18
|||||
Db 4 EAKICHOIEYFGDF 18

RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316.630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817.953

;; PRIOR FILING DATE: 1997-10-06
;; PRIOR APPLICATION NUMBER: 08/321,427
;; PRIOR FILING DATE: 1994-10-11
;; PRIOR APPLICATION NUMBER: 60/086,527
;; PRIOR FILING DATE: 1998-05-22
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
;; OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 85.3%; Score 87; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 EAKICHOIEYFGDF 18
| | | | | | | | | | | | | | | | | |
Db 4 EAKICHOIEYFGDF 18

RESULT 3
US-09-082-279B-1195
;; Sequence 1195, Application US/09082279B
;; Patent No. 6258782
;; GENERAL INFORMATION:
;; APPLICANT: Barney, Shawn
;; APPLICANT: Guthrie, Kelly
;; APPLICANT: Merutka, Gene
;; APPLICANT: Anwer, Maimed
;; APPLICANT: Lambert, Dennis
;; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
;; TITLE OF INVENTION: PHARMACOKINETIC PROPERTIES
;; FILE REFERENCE: 7872-043
;; CURRENT APPLICATION NUMBER: US/09/082,279B
;; CURRENT FILING DATE: 1998-05-20
;; NUMBER OF SEQ ID NOS: 1515
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1195
;; LENGTH: 24
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Core polypeptide
US-09-082-279B-1195

Query Match 34.3%; Score 35; DB 4; Length 24;
Best Local Similarity 46.2%; Pred. No. 22;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QOQEAKEKICHOIEY 13
| | | | | | | | | | | | | | | | | |
Db 2 QEWQEKIMHREDFY 14

RESULT 4
US-09-315-304B-1195
;; Sequence 1195, Application US/09315304B
;; Patent No. 6348568
;; GENERAL INFORMATION:
;; APPLICANT: Barney, S.
;; APPLICANT: Guthrie, K.
;; APPLICANT: Merutka, G.
;; APPLICANT: Anwer, M.
;; APPLICANT: Lambert, D.
;; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED PHARMACOKINETIC
;; TITLE OF INVENTION: PROPERTIES
;; FILE REFERENCE: 7872-052
;; CURRENT APPLICATION NUMBER: US/09/315,304B

;; CURRENT FILING DATE: 1999-05-20
;; PRIOR APPLICATION NUMBER: 09/082,279
;; PRIOR FILING DATE: 1998-05-20
;; NUMBER OF SEQ ID NOS: 1667
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1195
;; LENGTH: 24
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Core polypeptide
US-09-315-304B-1195

Query Match 34.3%; Score 35; DB 4; Length 24;
Best Local Similarity 46.2%; Pred. No. 22;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QOQEAKEKICHOIEY 13
| | | | | | | | | | | | | | | | | |
Db 2 QEWQEKIMHREDFY 14

RESULT 5
US-08-166-930-15
;; Sequence 15, Application US/08166930
;; Patent No. 5595678
;; GENERAL INFORMATION:
;; APPLICANT: Kraus, Michael
;; APPLICANT: Stuber, Werner
;; TITLE OF INVENTION: Synthetic Peptides, Antibodies Against
;; TITLE OF INVENTION: Them and Their Use
;; NUMBER OF SEQUENCES: 19
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
;; ADDRESSEE: Dunner
;; STREET: 1300 I Street, N.W., Suite 700
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20005-3315
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/166,930
;; FILING DATE: 15-DEC-1993
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: DE P 42 42 736.3
;; FILING DATE: 17-DEC-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Einaudi, Carol P.
;; REGISTRATION NUMBER: 32,220
;; REFERENCE/DOCKET NUMBER: 02481.1351-00000
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-408-4400
;; TELEFAX: 202-408-4400
;; INFORMATION FOR SEQ ID NO: 15:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 7 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-166-930-15

Query Match 29.4%; Score 30; DB 1; Length 7;
Best Local Similarity 80.0%; Pred. No. 2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 CHQIE 12
| | | | |

Db 1 CHQVE 5

RESULT 6

US-08-727-045A-15
; Sequence 15, Application US/08727045A
; Patent No. 5981697
; GENERAL INFORMATION:
; APPLICANT: Kraus, Michael
; APPLICANT: Stuber, Werner
; TITLE OF INVENTION: Synthetic Peptides, Antibodies Against
; TITLE OF INVENTION: Them and Their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: DC

ZIP: 20005-3315

COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/727,045A

FILING DATE: 08-OCT-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P 42 42 736.3

FILING DATE: 17-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Einaudi, Carol P.

REGISTRATION NUMBER: 32,220

REFERENCE/DOCKET NUMBER: 05552.1351-01000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4400

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-727-045A-15

Query Match 29.4%; Score 30; DB 2; Length 7;
Best Local Similarity 80.0%; Pred. NO. 2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQIE 12

Db 1 CHQVE 5

RESULT 7

US-09-408-172-15
; Sequence 15, Application US/09408172
; Patent No. 6441141
; GENERAL INFORMATION:
; APPLICANT: Kraus, Michael
; APPLICANT: Stuber, Werner
; TITLE OF INVENTION: Synthetic Peptides, Antibodies Against
; TITLE OF INVENTION: Them and Their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: DC

COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/408,172
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/727,045
FILING DATE: 08-OCT-1996
APPLICATION NUMBER: DE P 42 42 736.3
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Einaudi, Carol P.
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 05552.1351-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-408-172-15

Query Match 29.4%; Score 30; DB 4; Length 7;
Best Local Similarity 80.0%; Pred. NO. 2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQIE 12

Db 1 CHQVE 5

RESULT 8

US-08-618-696-20
; Sequence 20, Application US/08618696
; Patent No. 5861475
; GENERAL INFORMATION:
; APPLICANT: COOPER, JR., J. ALLEN D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/618,696
FILING DATE: 20-MAR-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/995,269
FILING DATE: 12/21/92
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UOAB:002/PAR
TELECOMMUNICATION INFORMATION:

TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-618-696-20

Query Match 29.4%; Score 30; DB 2; Length 11;
Best Local Similarity 44.4%; Pred. No. 62;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 HQIEYFGD 17
DB 1 NEASYFFGD 9

RESULT 9
US-08-211-312-9
; Sequence 9, Application US/08211312
; Patent No. 5986051
; GENERAL INFORMATION:
; APPLICANT: LABIGNE, AGNES
; APPLICANT: CUSSAC, VALERIE
; APPLICANT: FERRERO, RICHARD
; TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
; TITLE OF INVENTION: FOR THE REGULATION AND MATURATION OF UREASE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/211.312
APPLICATION NUMBER: US/08/211.312
FILING DATE: 01-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: FR 91 12198
FILING DATE: 03-OCT-1991

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR92/00921
FILING DATE: 02-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5986051man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-075-0XPCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-211-312-9

Query Match 29.4%; Score 30; DB 2; Length 11;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 AKICHOI 11
DB 1 AKICYEI 7

RESULT 10
US-09-033-753-20
; Sequence 20, Application US/09033753
; Patent No. 6017883
; GENERAL INFORMATION:
; APPLICANT: COOPER, JR., J. ALLEN D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,753
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/618,696
FILING DATE: 20-MAR-1996
APPLICATION NUMBER: 07/995,269
FILING DATE: 12/21/92
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: USAB:002/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-033-753-20

Query Match 29.4%; Score 30; DB 3; Length 11;
Best Local Similarity 44.4%; Pred. No. 62;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 HQIEYFGD 17
DB 1 NEASYFFGD 9

RESULT 11
US-08-472-285-9
; Sequence 9, Application US/08472285
; Patent No. 6027878
; GENERAL INFORMATION:
; APPLICANT: LABIGNE, AGNES
; APPLICANT: CUSSAC, VALERIE
; APPLICANT: FERRERO, RICHARD
; TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
; TITLE OF INVENTION: FOR THE REGULATION AND MATURATION OF UREASE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C. Jefferson Davis Highway, Suite 400
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,285
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/211,312
FILING DATE: 01-JUL-1994
APPLICATION NUMBER: FR 91 12198
FILING DATE: 03-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR92/00921
FILING DATE: 02-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, No. 6027878man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-075-0XPCT
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-472-285-9

Query Match 29.4%; Score 30; DB 3; Length 11;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 AKICHQI 11
| | | | |
Db 1 AKICYEI 7

RESULT 12
US-08-472-929-9
; Sequence 9, Application US/08472929
; Patent No. 6271017
; GENERAL INFORMATION:
; APPLICANT: LABIGNE, AGNES
; APPLICANT: CUSSAC, VALERIE
; APPLICANT: FERRERO, RICHARD
; TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
; TITLE OF INVENTION: FOR THE REGULATION AND MATURATION OF UREASE AND THEIR USE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,929
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/211,312
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR92/00921
FILING DATE: 02-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, No. 6271017man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-075-0XPCT
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-472-929-9

Query Match 29.4%; Score 30; DB 4; Length 11;
Best Local Similarity 71.4%; Pred. No. 62;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 AKICHQI 11
| | | | |
Db 1 AKICYEI 7

RESULT 13
US-08-618-696-11
; Sequence 11, Application US/08618696
; Patent No. 5861475
; GENERAL INFORMATION:
; APPLICANT: COOPER, JR., J. ALLEN D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,696
; FILING DATE: 20-MAR-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/995,269
; FILING DATE: 12/21/92
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UOAB:002/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-320-7200
; TELEFAX: 512-474-7577
; TELEX: NOT APPLICABLE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acid residues
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-618-696-11

Query Match 28.4%; Score 29; DB 2; Length 10;
 Best Local Similarity 50.0%; Pred. No. 82;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYFGD 17
 : :
 Db 1 EASYFFGD 8

RESULT 14

US-09-033-753-11
 ; Sequence 11, Application US/09033753
 ; Patent No. 6017883

GENERAL INFORMATION:

APPLICANT: COOPER, JR., J. ALLEN D.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
 INHIBITION OF PHAGOCYTES
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ARNOLD, WHITE & DURKEE
 STREET: P.O. BOX 4433
 CITY: HOUSTON
 STATE: TEXAS
 COUNTRY: USA
 ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/033,753
 FILING DATE: 12/21/92
 CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/618,696
 FILING DATE: 20-MAR-1996
 APPLICATION NUMBER: 07/995,269
 FILING DATE: 12/21/92
 ATTORNEY/AGENT INFORMATION:
 NAME: PARKER, DAVID L.
 REGISTRATION NUMBER: 32,165
 REFERENCE/DOCKET NUMBER: UOAB:002/PAR

TELECOMMUNICATION INFORMATION:

TELEPHONE: 512-320-7200
 TELEFAX: 512-474-7577
 TELEX: NOT APPLICABLE
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acid residues
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-09-033-753-11

Query Match 28.4%; Score 29; DB 3; Length 10;
 Best Local Similarity 50.0%; Pred. No. 82;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYFGD 17
 : :
 Db 1 EASYFFGD 8

RESULT 15

US-08-618-696-7
 ; Sequence 7, Application US/08618696

Patent No. 5861475
 GENERAL INFORMATION:
 APPLICANT: COOPER, JR., J. ALLEN D.
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
 INHIBITION OF PHAGOCYTES
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ARNOLD, WHITE & DURKEE
 STREET: P.O. BOX 4433
 CITY: HOUSTON
 STATE: TEXAS
 COUNTRY: USA
 ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/618,696
 FILING DATE: 20-MAR-1996
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/995,269
 FILING DATE: 12/21/92
 ATTORNEY/AGENT INFORMATION:

NAME: PARKER, DAVID L.
 REGISTRATION NUMBER: 32,165
 REFERENCE/DOCKET NUMBER: UOAB:002/PAR
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 512-320-7200
 TELEFAX: 512-474-7577
 TELEX: NOT APPLICABLE
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 11 amino acid residues
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-618-696-7

Query Match 28.4%; Score 29; DB 2; Length 11;
 Best Local Similarity 50.0%; Pred. No. 91;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYFGD 17
 : :
 Db 2 EASYFFGD 9

Search completed: April 23, 2003, 13:50:25
 Job time : 8.49438 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:28:18 ; Search time 27.2022 Seconds
(without alignments)
88.173 Million cell updates/sec

Title: US-09-836-073-4
Perfect score: 102
Sequence: 1 QQQEAKICHOIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 320064

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.1
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	87	85.3	18	AAV52200	Human la autoantigen
2	73	71.6	21	AAK43394	La/Ssb epitope 17.
3	37	36.3	10	ABH84046	Transglutaminase 1
4	36	35.3	23	AAU00924	VH ligand-binding
5	35	34.3	24	AAV89809	Core polypeptide f
6	35	34.3	24	ABO1217	Viral DP178/107-11
7	35	34.3	24	ABO2668	Viral core polypep
8	35	34.3	24	AAU13763	DP178-like/DP107-1
9	35	34.3	24	AAV78210	Core polypeptide T
10	31	30.4	8	AAK43395	La/Ssb epitope 24.

11	31	30.4	9	20	AAV10410	T cell epitope/MHC
12	31	30.4	10	23	AAE22202	Murine MC-1 antio
13	31	30.4	11	20	AAW52555	Anti-progesterone
14	31	30.4	16	22	AAW89909	Vaccine related MH
15	31	30.4	18	20	AAW93362	Human DPD protein
16	31	30.4	19	18	AAW27677	Coagulation inhibi
17	31	30.4	23	23	AAU99805	Biomembrane perma
18	30.5	29.9	23	20	AAV25901	Human secreted pro
19	30	29.4	7	15	AAV56273	Peptide antigen us
20	30	29.4	11	14	AAK34400	Fragment of Helico
21	30	29.4	11	15	AAV56307	Modified Influenza
22	30	29.4	15	18	AAK36481	VAF1 homologous pe
23	30	29.4	21	22	ABK48862	Mutant human insul
24	30	29.4	21	23	ABG66356	IgE Fcpsiilon RI b
25	30	29.4	22	20	AAV19621	SEQ ID NO 339 from
26	30	29.4	25	20	AAV36636	Fragment of human
27	29	28.4	9	23	AAU92267	PHORI-F5D6 peptide
28	29	28.4	9	23	AAU92288	PHORI-F5D6 peptide
29	29	28.4	9	23	AAU92398	PHORI-F5D6 peptide
30	29	28.4	9	23	AAU92773	PHORI-F5D6 peptide
31	29	28.4	9	23	AAU92875	PHORI-F5D6 peptide
32	29	28.4	10	15	AAV56297	Synthetic modified
33	29	28.4	10	23	AAU92318	PHORI-F5D6 peptide
34	29	28.4	10	23	AAU92332	PHORI-F5D6 peptide
35	29	28.4	10	23	AAU92548	PHORI-F5D6 peptide
36	29	28.4	10	23	AAU92639	PHORI-F5D6 peptide
37	29	28.4	10	23	AAU92744	PHORI-F5D6 peptide
38	29	28.4	10	23	AAU92821	PHORI-F5D6 peptide
39	29	28.4	10	23	AAU92931	PHORI-F5D6 peptide
40	29	28.4	10	23	AAU92942	PHORI-F5D6 peptide
41	29	28.4	11	15	AAV56292	Synthetic derivati
42	29	28.4	12	21	AAV93796	Reactive peptide w
43	29	28.4	12	21	AAV93875	S. acidocaldarius
44	29	28.4	14	17	AAK90665	Hepatitis GB virus
45	29	28.4	14	21	AAK909328	

ALIGNMENTS

RESULT 1
AAV52200
ID AAV52200 standard; peptide: 18 AA.

XX AC AAV52200;
XX DT 14-MAR-2000 (first entry)
XX DE Human la autoantigen peptide (LAP).

XX KW La autoantigen; LAP; internal ribosome entry site; IRES; translation;
XX KW viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
XX KW coronavirus; hepatitis virus; rhinovirus; adenovirus; coxsackie virus;
XX KW parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
XX KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
XX KW vesicular stomatitis virus.

XX OS Homo sapiens.

XX PN WO9961613-A2.

XX PD 02-DEC-1999.

XX PF 21-MAY-1999; 99WO-US11281.

XX PR 22-MAY-1998; 98US-0086527.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Das S, Dasgupta A;

XX DR WPI; 2000-062712/05.

XX XX

PT New yeast inhibitory peptide useful for inhibiting viral protein
 PT translation and replication -
 XX Claim 5; Page 57; 81pp; English.
 XX This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (IRNA see A445200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rhinovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.
 XX

Sequence 18 AA; Query Match 85.3%; Score 87; DB 21; Length 18;
 Best Local Similarity 100.0%; Pred. No. 2e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4' EAKICHQIEYYFGDF 18
 DB 4 EAKICHQIEYYFGDF 18
 |||||

RESULT 2
 AAR43394
 ID AAR43394 standard; peptide: 21 AA.
 AC AAR43394;
 DT 12-MAY-1994 (first entry)
 DE La/SSB epitope 17.

Linear: epitope: 60 kD; Ro/SSA: La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

OS Homo sapiens.
 XX WO9321223-A.
 XX 28-OCT-1993.
 XX 13-APR-1993; 93WO-US03484.
 XX 13-APR-1992; 92US-0867819.
 XX (OKLA) UNIV OKLAHOMA STATE.
 XX Harley JB;
 XX WPI; 1993-351658/44.
 XX

New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribo:nucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus
 XX

Claim 1; Page 30; 43pp; English.
 XX The sequences given in AAR43391-562 are linear epitopes which are

CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B',
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polypyridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.
 XX

Sequence 21 AA;

Query Match 71.6%; Score 73; DB 14; Length 21;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICHQIEYYFGDF 18
 DB 1 ICHQIEYYFGDF 12
 |||||

RESULT 3
 ABB84046
 ID ABB84046 standard; peptide: 10 AA.
 AC ABB84046;
 DT 21-AUG-2002 (first entry)
 DE Transglutaminase inhibitory peptide cr type #16.

Transglutaminase inhibitor; Factor XIII inhibitor; XIIIa inhibitor;
 KW ophthalmological; antiinflammatory; antirheumatic; antiarthritic;
 KW thrombolytic; neuroprotective; nootropic; antiseborrheic; dermatological;
 KW cytostatic; anti-HIV; antipsoriatic; cataract; inflammatory disease;
 KW arthritis; thrombosis; Alzheimer's disease; Huntington's chorea;
 KW acne; cancer; HIV infection; psoriasis.
 XX

Unidentified.

WO200236798-A2.
 XX 10-MAY-2002.
 XX 02-NOV-2001; 2001WO-EPI2727.
 XX 03-NOV-2000; 2000DE-1054687.
 XX (NZYM-) N ZYME BIOTEC GMBH.
 XX Fuchsbauser H, Pasternack R, Zotzel J;
 XX WPI; 2002-444364/47.

New amino acid or peptide derivatives or analogs, are selective
 PT transglutaminase inhibitors useful e.g. for treating cataract,
 PT inflammatory diseases, rheumatoid arthritis, thrombosis, Alzheimer's
 PT disease and cancer
 XX

Disclosure; Page 13; 44pp; German.

This invention describes novel amino acid or peptide derivatives or
 CC analogues (I), containing a modified side-chain (e.g. containing a formyl

CC group) which are transglutaminase inhibitors and Factor XIII/XIIIa
 CC inhibitors. The products of the invention have ophthalmological,
 CC anti-inflammatory, antirheumatic, antiarthritic, thrombolytic,
 CC neuroprotective, nootropic, antiseborrheic, dermatological, cytostatic,
 CC anti-HIV and antiproliferative activity. (I) are transglutaminase inhibitors,
 CC especially inhibitors of crosslinking of proteins or peptides
 CC (specifically fibrin and/or alpha₂-plasmin inhibitor), incorporation of
 CC primary amines in proteins and peptides, hydrolysis of the
 CC gamma-carboxamide group of glutamine residues bound in proteins or
 CC peptides, blood factor XIII/XIIIa and mammalian, human, tissue, liver,
 CC brain, eye lens, keratinocyte, epidermal, prostate, plant, parasitic
 CC and/or bacterial transglutaminases. The products of the invention can be
 CC used for treating cataract, inflammatory diseases, rheumatoid arthritis,
 CC chronic arthritis, thrombosis, Alzheimer's disease, Huntington's chorea,
 CC acne, cancer (by induction of apoptosis), HIV infections and psoriasis.
 CC (I) are targeted and specific transglutaminase inhibitors, which can
 CC inhibit a specific type of transglutaminase in the human or animal body
 CC without affecting other transglutaminases. ABB84001-ABB84049 represent
 CC transglutaminase inhibitors described in the method of the invention.
 XX
 SQ Sequence 10 AA;

Query Match 36.3%; Score 37; DB 23; Length 10;
 Best Local Similarity 66.7%; Pred. No. 16;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 2 QQEKICHQ 10
 I::I I I I I
 DB 1 QKQAPICHQ 9

RESULT 4
 AAU00924
 ID AAU00924 standard; Peptide; 23 AA.
 XX
 AC AAU00924;
 XX
 DT 18-MAY-2001 (first entry)
 XX
 DE VH ligand-binding domain variant CDR3 region #31.
 XX
 KW Human; monoclonal antibody A6; VH ligand-binding domain; variable region;
 KW parental ligand binding molecule; PLBM; framework region; FR; CDR;
 KW combinatorial library; diagnosis; therapy; surface protein; pesticide;
 KW cell-receptor; carbohydrate; lipid; antibiotic; steroid; hormone; dye;
 KW herbicide; polymer; immunoglobulin; CDR3 region; M2 immunoglobulin;
 KW panning.
 XX
 OS Homo sapiens.
 XX
 PN WO200118058-A2.
 XX
 PD 15-MAR-2001.
 XX
 PF 07-SEP-2000; 2000WO-CA01027.
 XX
 PR 07-SEP-1999; 99CA-2282179.
 PR 04-NOV-1999; 99US-0163546.
 XX
 PA (NOVO-) NOVOPHARM BIOTECH INC.
 XX
 PI Kaplan H, Entwistle J, Tanha J, Narang S, Dan M;
 XX
 DR WPI; 2001-235191/24.
 XX
 PT Combinatorial libraries including phage display library comprises
 PT variants of immunoglobulin VH fragments which comprises the framework
 PT regions of wild-type or modified immunoglobulin VH domain of human A6
 PT antibody -
 XX
 PS Disclosure; Page 25; 133pp; English.
 XX
 CC The sequence represents a variant CDR3 region of a parental VH

CC ligand-binding domain of human monoclonal antibody A6, isolated by
 CC panning against M2 immunoglobulin. Variants of a parental ligand
 CC binding molecule (PLBM) comprise an immunoglobulin (Ig) VH fragment
 CC including the framework (FR) regions of the wild-type/modified Ig VH
 CC domain of human monoclonal antibody A6. The variant differs from PLBM at
 CC residues constituting part of one of the CDRs of PLBM. A library
 CC expressing binding domains can be created by cloning a parental DNA
 CC sequence that encodes a parental domain, in order to produce parental
 CC clones, and replacing a variable region of the clones with a variant DNA
 CC sequence. A library of genetic packages can then be generated, each
 CC package having a surface and a surface protein binding domain expressed
 CC by the variant DNA. Combinatorial libraries are useful for generating
 CC forms of recombinant antibody fragments that specifically bind target
 CC ligands including natural and synthetic molecules and macromolecules to
 CC be used in diagnostic and/or therapeutic processes. Examples include
 CC enzymes, cell-receptors, carbohydrates, lipids, and organic-based and
 CC inorganic-based molecules such as antibiotics, steroids, hormones,
 CC pesticides, herbicides, dyes and polymers.
 XX

SQ Sequence 23 AA;

Query Match 35.3%; Score 36; DB 22; Length 23;
 Best Local Similarity 45.5%; Pred. No. 58;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 8 CHQIEYFQDF 18
 I::I I I I I
 DB 11 CFMVDYKFSDF 21

RESULT 5
 AAY89809
 ID AAY89809 standard; peptide; 24 AA.
 XX
 AC AAY89809;
 XX
 DT 23-MAY-2000 (first entry)
 XX
 DE Core polypeptide fragment T No. 1377.
 XX
 KW Retrovirus; hybrid polypeptide; enhancer; gp41; envelope protein; HIV-1;
 KW HIV-2; SIV; pharmacokinetic; half-life; growth factor; cytokine; viral;
 KW anti-fusogenic; differentiation factor; interleukin; interferon;
 KW colony stimulating factor; hormone; angiogenic factor.
 XX
 OS Unidentified.
 XX
 PN WO9959615-A1.
 XX
 PD 25-NOV-1999.
 XX
 PF 20-MAY-1999; 99WO-US11219.
 XX
 PR 20-MAY-1998; 98US-0082279.
 XX
 PA (TRIM-) TRIMERIS INC.
 XX
 PI Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;
 XX
 DR WPI; 2000-136792/12.
 XX
 PT A new hybrid polypeptide with enhanced pharmacokinetic properties
 PT comprises enhancer sequence -
 XX
 PS Disclosure; Page 44; 124pp; English.
 XX
 CC The invention relates to hybrid polypeptides comprising enhancer peptide
 CC sequence linked to core polypeptides. The enhancer polypeptides are
 CC derived from various retroviral envelope (gp41) protein sequences,
 CC especially from HIV-1, HIV-2 and SIV. The enhancer peptides enhance the
 CC pharmacokinetic properties such as increasing the half-life of any core
 CC polypeptide that they are linked to. The core polypeptides are any
 CC polypeptide that may be introduced into a living system and that can

CC function as a pharmacologically useful peptide for the treatment or
 CC prevention of a disease. The core polypeptides are bioactive peptides
 CC selected from a growth factor, cytokine, differentiation factor,
 CC interleukin, interferon, colony stimulating factor, hormone or
 CC angiogenic factor. The peptides of the invention can be used for
 CC inhibiting viral infection and can be used in anti-viral and
 CC anti-fusogenic treatments. Sequences AAY8651-Y9055 represent core
 CC polypeptide fragments that can be used in the invention. Some sequences
 CC among those indicated also comprise enhancer fragments at terminal ends
 CC and form hybrid polypeptides.

XX SQ Sequence 24 AA;

Query Match 34.3%; Score 35; DB 21; Length 24;

Best Local Similarity 46.2%; Pred. No. 88;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 QQQEAKICHQIEY 13

DB 2 QWEQKIMHRFDY 14

RESULT 6

ABB01217

ID ABB01217 standard; Peptide; 24 AA.

XX AC ABB01217;

XX DT 03-JAN-2002 (first entry)

XX DE Viral DPI78/107-like region peptide T1377.

XX KW Human immunodeficiency virus; HIV; respiratory syncytial virus; RSV;
 XX KW virucide; heptad repeat region; transmembrane protein; gp41; HRI; HR2;
 XX KW infection.

XX OS Viridiae.

XX PH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal is substituted by Ac"

FT Modified-site 24

FT /note= "C-terminal amide"

XX WO200164013-A2.

XX PD 07-SEP-2001.

XX PF 07-FEB-2001; 2001WO-US03988.

XX PR 29-FEB-2000; 2000US-0515965.

XX PA (TRIM-) TRIMERIS INC.

XX PI Antczak JB, Delmedico MK, Erickson JB, Lambert DM, Sista P;

XX DR WPI; 2001-514829/56.

XX FT Heptad repeat region peptide analogs useful for inhibiting virus/cells
 XX FT fusion, useful for treating HIV and Respiratory Syncytial Virus
 XX FT infection.

XX PS Disclosure; Page 57; 587pp; English.

XX CC The invention relates to isolated analogues of the heptad repeat region
 XX CC peptides DPI78 and DPI107. DPI78 and DPI107 correspond to amino acids
 XX CC 638-673 (heptad repeat region HR2) and 558-595 (heptad repeat region
 XX CC HRI) respectively, of HIV-1LAI transmembrane protein gp41. The HRI
 XX CC and HR2 regions of proteins interact non-covalently with each other
 XX CC and/or with peptides derived from them. This interaction is required for
 XX CC normal infectivity of viruses such as RSV and HIV. The heptad
 XX CC repeat region peptide analogues may be used to inhibit respiratory
 XX CC syncytial virus (RSV) infection in a cell. They may also be used to

CC inhibit HIV infection. The present sequence is a peptide provided in
 CC the specification.

XX SQ Sequence 24 AA;

Query Match 34.3%; Score 35; DB 22; Length 24;

Best Local Similarity 46.2%; Pred. No. 88;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 QQQEAKICHQIEY 13

DB 2 QWEQKIMHRFDY 14

RESULT 7

ABB02668

ID ABB02668 standard; Peptide; 24 AA.

XX AC ABB02668;

XX DT 03-JAN-2002 (first entry)

XX DE Viral core polypeptide, SEQ ID NO: 1195.

XX KW Human immunodeficiency virus; HIV; respiratory syncytial virus; RSV;
 XX KW virucide; heptad repeat region; transmembrane protein; gp41; HRI; HR2;
 XX KW infection.

XX OS Viridiae.

XX PN WO200164013-A2.

XX PD 07-SEP-2001.

XX PF 07-FEB-2001; 2001WO-US03988.

XX PR 29-FEB-2000; 2000US-0515965.

XX PA (TRIM-) TRIMERIS INC.

XX PI Antczak JB, Delmedico MK, Erickson JB, Lambert DM, Sista P;

XX DR WPI; 2001-514829/56.

XX PT Heptad repeat region peptide analogs useful for inhibiting virus/cells
 XX PT fusion, useful for treating HIV and Respiratory Syncytial Virus
 XX PT infection.

XX PS Disclosure; Page 483; 587pp; English.

XX CC The invention relates to isolated analogues of the heptad repeat region
 XX CC peptides DPI78 and DPI107. DPI78 and DPI107 correspond to amino acids
 XX CC 638-673 (heptad repeat region HR2) and 558-595 (heptad repeat region
 XX CC HRI) respectively, of HIV-1LAI transmembrane protein gp41. The HRI
 XX CC and HR2 regions of proteins interact non-covalently with each other
 XX CC and/or with peptides derived from them. This interaction is required for
 XX CC normal infectivity of viruses such as RSV and HIV. The heptad
 XX CC repeat region peptide analogues may be used to inhibit respiratory
 XX CC syncytial virus (RSV) infection in a cell. They may also be used to
 XX CC inhibit HIV infection. The present sequence is a peptide provided in
 XX CC the specification.

XX SQ Sequence 24 AA;

Query Match 34.3%; Score 35; DB 22; Length 24;

Best Local Similarity 46.2%; Pred. No. 88;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 QQQEAKICHQIEY 13

DB 2 QWEQKIMHRFDY 14

XX DT 12-MAY-1994 (first entry)
 XX DE La/SSB epitope 24.
 XX KW Linear; epitope: 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
 XX OS Homo sapiens.
 XX PN WO9321223-A.
 XX PD 28-OCT-1993.
 XX PF 13-APR-1993; 93WO-US03484.
 XX PR 13-APR-1992; 92US-0867819.
 XX PA (OKLA) UNIV OKLAHOMA STATE.
 XX PI Harley JB;
 XX DR WPI; 1993-351658/44.
 XX PT New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus
 XX PS Claim 1; Page 30; 43pp; English.
 XX CC The sequences given in AAR43391-562 are linear epitopes which are
 CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polyuridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.
 XX SQ Sequence 8 AA;
 Query Match 30.4%; Score 31; DB 14; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 14 YFGDF 18
 |||||
 Db 1 YFGDF 5
 RESULT 11
 AAY10410
 ID AAY10410 standard; Peptide; 9 AA.
 XX AC AAY10410;
 XX DT 12-MAY-1999 (first entry)
 XX KW

DE T cell epitope/MHC ligand SEQ ID NO:340.
 XX Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;
 KW immunisation; tumour; infectious disease; immunotherapy; cancer;
 KW malignant melanoma; viral disease; hepatitis; AIDS.
 OS Synthetic.
 OS Simian virus.
 XX WO9902183-A2.
 XX 21-JAN-1999.
 XX 10-JUL-1998; 98WO-US14289.
 XX 10-DEC-1997; 97US-0988320.
 XX 10-JUL-1997; 97CA-2209815.
 XX (CTLI-) CTL IMMUNOTHERAPIES CORP.
 XX Kuendig TM, Simard JUL;
 PI WPI; 1999-120514/10.
 DR Inducing a cytotoxic T lymphocyte response - by maintaining a level
 of antigen in the lymphatic system of a mammal so as to provide a
 sustained CTL response, used to treat, e.g. AIDS
 PS Disclosure; Page 37; 199pp; English.
 XX CC The present invention describes a method of inducing and/or sustaining
 CC an immunological cytotoxic T lymphocyte (CTL) response in a mammal. The
 CC method comprises: (a) delivering an antigen to the mammal at a level to
 CC induce an immunological CTL response in the mammal; and (b) maintaining
 CC the level of the antigen in the mammal's lymphatic system to maintain
 CC the immunologic CTL response. The method can be used for the delivery of
 CC e.g. a differentiation antigen, a tumour-specific multilineage antigen,
 CC an embryonic antigen, an oncogene antigen, a mutated tumour-suppressor
 CC gene antigen, or a viral antigen. They can be used for the treatment of
 CC disease such as cancer, e.g. malignant melanoma or infectious disease,
 CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery
 CC to the lymphatic system provides for potent CTL stimulation that takes
 CC place in the milieu of the lymphoid organ, and it sustains stimulation
 CC that is necessary to keep CTL active, cytotoxic and recirculating
 CC through the body. AAY10071 to AAY10639 represent examples of peptide
 CC antigens given in the present invention.
 XX SQ Sequence 9 AA;
 Query Match 30.4%; Score 31; DB 20; Length 9;
 Best Local Similarity 75.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 4 EAKICHQI 11
 |||||
 Db 2 EAEIAHQI 9
 RESULT 12
 AAE22202
 ID AAE22202 standard; peptide; 10 AA.
 XX AC AAE22202;
 XX 25-JUL-2002 (first entry)
 XX Murine MC-1 antibody heavy chain variable domain (VH) CDR3 peptide.
 XX Chemokine construct; human immunodeficiency virus 1; allergic disease;
 KW skin disease; immunological disorder; autoimmune disease; psoriasis;
 KW multiple sclerosis; rheumatoid arthritis; inflammatory bowel disease;
 KW diabetes; skin inflammation; atopic dermatitis; inflammatory disease;
 KW inflammatory renal disease; HIV-1; transplant rejection; murine; CDR3;

KW heavy chain variable domain; VH; complementarity determining region 3;
 KW MC-1; antibody.
 XX
 OS Mus sp.
 XX
 PN WO200220615-A2.
 XX
 PD 14-MAR-2002.
 XX
 XX 10-SEP-2001; 2001WO-EPI0433.
 PF
 XX 08-SEP-2000; 2000EP-0119694.
 PR
 PR 05-SEP-2001; 2001US-0948004.
 XX
 XX (MYCR-) MICROMET AG.
 PA
 XX Mack M, Schloendorff D, Spring M;
 PI
 XX WPI; 2002-362240/39.
 DR
 XX Use of an antibody and/or chemokine construct that binds to a chemokine
 PT receptor, for eliminating cells latently infected with primate
 PT immunodeficiency virus, or treating, preventing and alleviating immune
 PT disorders.
 XX
 PS Disclosure; Page 117; 117pp; English.
 XX
 CC The invention relates to the use of an antibody and/or chemokine
 CC construct that binds a chemokine receptor for preparing a pharmaceutical
 CC composition for eliminating cells latently infected with a primate
 CC immunodeficiency virus (e.g. human immunodeficiency virus (HIV-1)). They
 CC are used in gene therapy and as vaccines. The antibody and/or chemokine
 CC construct is also used for preparing a pharmaceutical composition for
 CC treating, preventing and/or alleviating immunological disorders
 CC including autoimmune diseases (e.g. multiple sclerosis, type I diabetes
 CC and rheumatoid arthritis), allergic diseases, skin diseases (e.g. skin
 CC inflammation, atopic dermatitis and psoriasis), inflammatory diseases
 CC such as inflammatory joint diseases (chronic arthritis), inflammatory
 CC renal diseases and inflammatory bowel diseases and graft versus host
 CC and transplant rejections. The present sequence is murine MC-1 antibody
 CC heavy chain variable domain (VH) complementarity determining region 3
 CC (CDR3) peptide.
 XX
 SQ Sequence 10 AA;
 Query Match 30.4%; Score 31; DB 23; Length 10;
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 12 EYFGDF 18
 Db |||:|
 2 EYFGDF 8
 RESULT 13
 AAW95255
 ID AAW95255 standard; peptide; 11 AA.
 XX
 AC AAW95255;
 XX
 XX 11-MAR-1999 (first entry)
 DT
 DE Anti-progesterone antibody Vh region (clone 1578/p2) CDR H3 sequence.
 DE
 KW ARM; protein selection; display; cell free system; genetic information;
 KW reverse transcription; single-chain antibody; RT-PCR; primer; drug;
 KW antibody-ribosome-mRNA complex; antibody engineering; progesterone;
 KW testosterone; human
 XX
 OS Homo sapiens.
 XX
 PN WO9854312-A1.
 XX

PD 03-DEC-1998.
 XX
 XX 28-MAY-1998; 98WO-GB01564.
 PF
 XX 28-FEB-1998; 98GB-0004195.
 PR
 PR 28-MAY-1997; 97GB-0010829.
 XX
 XX 26-NOV-1997; 97GB-0024850.
 XX
 PA (BADR-) BABRAHAM INST.
 XX
 XX He M, Taussig MJ;
 PI
 XX WPI; 1999-059832/05.
 DR
 XX In vitro display and evolution of proteins - by transcription and
 PT translation of DNA in a cell free system, selection and recovery of
 PT complexes and RT-PCR on RNA bound to ribosomes
 XX
 PS Example 13; Fig 19; 62pp; English.
 XX
 CC The invention relates to methods for the display and selection of
 CC proteins or peptides and for recovery of the genetic material encoding
 CC them. One method comprises (a) transcription and translation of DNA in a
 CC cell free system such that complexed particles are formed, each
 CC comprising at least one individual nascent protein or peptide or other
 CC DNA expression product associated with one or more ribosomes and the
 CC specific mRNA encoding the protein or peptide; (b) contacting the
 CC complexed particles with a ligand, antigen, antibody or other agent in
 CC order to select particles through binding to the protein or peptide
 CC product; and (c) recovering the genetic information encoding the protein
 CC or peptide as DNA by RT-PCR carried out on the mRNA while the latter
 CC remains bound to the complexed particle. The steps of display, selection
 CC and recovery can be repeated in consecutive cycles. The method is
 CC exemplified using single-chain antibody constructs as antibody-ribosome-
 CC mRNA complexes (ARMS). Methods in which the DNA is produced by RT-PCR,
 CC methods for making antibodies of human, mouse or rat are also provided.
 CC The methods can be used for the display and selection of single chain
 CC antibody fragments from libraries, antibody engineering, selection of
 CC human antibodies and selection of proteins from mRNA libraries. They can
 CC also be used to select ligands for combining sites or receptors, such
 CC ligands having potential uses as drugs or therapies. By carrying out the
 CC RT-PCR recovery step directly on the intact ribosome complex without
 CC prior dissociation to release the mRNA maximal efficiency and
 CC sensitivity can be obtained. Peptides AAW95247 to AAW95271 represent
 CC sequences of human anti-progesterone and anti-testosterone antibodies
 CC isolated from an immunised transgenic mouse by ARM display.
 XX
 SQ Sequence 11 AA;
 Query Match 30.4%; Score 31; DB 20; Length 11;
 Best Local Similarity 50.0%; Pred. No. 1.7e+02;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 QY 9 HQIEYFG 16
 Db ::||:|
 3 YEIDWYFG 10
 RESULT 14
 AAW98909
 ID AAW98909 standard; Peptide; 16 AA.
 XX
 AC AAW98909;
 XX
 XX 07-DEC-2001 (first entry)
 DT
 XX Vaccine related MHC ligand peptide SEQ ID NO:12.
 DE
 XX Glutamic acid; glutamine; vaccine; major histocompatibility complex;
 KW MHC; immunomodulator; antiallergic; endocrine; neuroprotectant;
 KW virucidal; bactericidal; antiparasitic; fungicidal; cytostatic;
 KW medicine; pharmaceutical; immune disorder; immune deficiency;
 KW autoimmune; hypersensitivity; allergy; graft rejection; infection;
 XX

KW hormonal disorder; central nervous system disease; cancer; melanoma;
 KW anti-melanoma vaccine; human immunodeficiency virus.

XX Rabies virus.

XX WO20010772-A2.

XX 27-SEP-2001.

XX 22-MAR-2001; 2001WO-FR00872.

XX 23-MAR-2000; 2000FR-0003711.

XX (FABR) FABRE MEDICAMENT SA PIERRE.

XX Klinguer-Hamou C, Corvaia N, Beck A, Goetsch L;

XX WPI; 2001-611470/70.

XX Stabilized pharmaceutical containing N-terminal glutamic acid or
 PT glutamine, useful e.g. in anti-melanoma vaccines, is an addition salt
 PT with strong acid

XX Claim 9; Page 31; 149pp; French.

XX The present invention describes a pharmaceutical compound (I) that
 CC contains an N-terminal glutamic acid (Glu) or glutamine (Gln) residue
 CC in the form of an addition salt with a strong, physiologically
 CC acceptable acid (II). Also described are: (a) a pharmaceutical
 CC composition containing at least one (I); (b) a vaccine containing at
 CC least one (I) where this is a major histocompatibility complex (MHC)
 CC ligand (Ia); (c) a method for in vitro diagnosis of diseases associated
 CC with the presence of (Ia); (d) a kit for method (c) that includes a (Ia);
 CC and (e) a process for preparing (I). (I) has immunomodulator, endocrine,
 CC antiallergic, neuroprotectant, virucidal, bactericidal, antiparasitic,
 CC fungicidal and cytostatic activities. (I) are useful, in human or
 CC veterinary medicine, in pharmaceutical compositions (for treating immune
 CC disorders, e.g. immune deficiency, autoimmune states, hypersensitivity,
 CC allergy, graft rejection, infection, hormonal disorders and central
 CC nervous system diseases), also, where (I) is a MHC ligand (Ia), in
 CC vaccines for treatment or prevention of: (i) viral, bacterial, parasitic
 CC or fungal infections; or (ii) of cancers. A particular application is in
 CC anti-melanoma vaccines. (I) are also useful for in vitro diagnosis of
 CC diseases associated with interactions between MHC and (I), e.g. melanoma
 CC and human immunodeficiency virus infection. AAM98898 to AAM99592
 CC represent peptides which can be used in pharmaceutical compounds from
 CC the present invention.

XX Sequence 16 AA;

Query Match 30.4%; Score 31; DB 22; Length 16;

Best Local Similarity 75.0%; Pred. No. 2.6e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHQI 11

DB 8 EAEIAHQI 15

RESULT 15

AAW93362

ID AAW93362 standard; Protein; 18 AA.

XX AAW93362;

XX 28-MAY-1999 (first entry)

XX Human DPD protein fragment #1.

XX DPD; dihydropyrimidine-dehydrogenase; monoclonal antibody; MAB;
 KW immunoassay reagent; cancer patient; treatment; antitumor agent;
 KW 5-fluorouracil; affinity purification; toxicity.

OS Homo sapiens.

XX DE19837391-A1.

XX 25-FEB-1999.

XX 18-AUG-1998; 98DE-1037391.

XX 22-AUG-1997; 97EP-0114630.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Hasegawa M, Yoshikubo T;

XX WPI; 1999-155202/14.

XX Monoclonal antibody specific for dihydropyrimidine dehydrogenase -
 PT for assessing patient response to 5-fluorouracil antitumor agents

XX Disclosure; Page 24; 34pp; German.

XX This invention describes a monoclonal antibody (MAB) specific for
 CC dehydropyrimidine dehydrogenase (DPD). This MAB is used as immunoassay
 CC reagents to identify a lack of DPD in a patient and to assess the
 CC sensitivity of cancer patients to treatment with antitumor agents of the
 CC 5-fluorouracil (5-FU) type. The MAB can also be used for affinity
 CC purification of DPD. DPD is involved in reduction of 5-FU (and related
 CC catabolites and derivatives) and lack of it is associated with increased
 CC toxicity of this type of antitumor agent. It has specific binding
 CC interaction. The MAB provide a sensitive and reliable test for DPD,
 CC which is simple, rapid and suitable for routine screening.

XX Sequence 18 AA;

Query Match 30.4%; Score 31; DB 20; Length 18;

Best Local Similarity 50.0%; Pred. No. 2.9e+02;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 8 CHOIEYFGD 17

DB 1 CEKLENNFGD 10

Search completed: April 23, 2003, 13:42:57

Job time : 28.2022 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:21:27 ; Search time 27.9101 Seconds
(without alignments)
132.885 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98
Sequence: 1 AALQAKICHQIYVFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	83	84.7	381	11 Q9CYB9	Q9cyb9 mus musculus
2	62	63.3	206	13 Q8QH15	Q8qh15 gallus gall
3	47	48.0	671	4 Q9UL65	Q9ul65 homo sapien
4	46	46.9	391	2 P72382	P72382 staphylococ
5	46	46.9	391	2 P95709	P95709 staphylococ
6	46	46.9	391	16 Q99X57	Q99x57 staphylococ
7	45	46.9	608	11 Q9QUG9	Q9qug9 mus musculus
8	45	45.9	568	16 Q9KKL7	Q9kk17 vibrio chol
9	45	45.9	928	10 Q9LJ02	Q9lj02 oryza sativ
10	44	44.9	239	10 Q9SMD8	Q9smd8 laminaria d
11	43	43.9	337	2 Q8RK98	Q8rk98 mycoplasma
12	42.5	43.4	577	16 Q97DB6	Q97db6 clostridium
13	42	42.9	396	5 O01806	O01806 caenorhabdi
14	42	42.9	541	16 Q92KY5	Q92ky5 helicobacte
15	42	42.9	542	16 O25534	O25534 helicobacte
16	42	42.9	658	10 Q9CAN3	Q9can3 arabidopsis

17	42	42.9	1220	12 Q39272	Q39272 equine herp
18	42	42.9	1249	12 Q65152	Q65152 african swi
19	42	42.9	2644	3 Q13535	Q13535 homo sapien
20	41	41.8	259	3 Q9P5X3	Q9p5x3 neurospora
21	41	41.8	296	5 Q9W129	Q9w129 drosophila
22	41	41.8	336	2 Q50359	Q50359 mycoplasma
23	41	41.8	336	16 Q8QD2	Q8qgd2 mycoplasma
24	41	41.8	366	2 Q85797	Q85797 mycoplasma
25	41	41.8	369	2 Q30382	Q30382 mycoplasma
26	41	41.8	369	2 Q9R8C2	Q9r8c2 mycoplasma
27	41	41.8	369	16 Q98PP1	Q98pp1 mycoplasma
28	41	41.8	399	2 Q85799	Q85799 mycoplasma
29	41	41.8	411	10 Q9FL36	Q9fl36 arabidopsis
30	41	41.8	422	10 Q94A38	Q94a38 arabidopsis
31	41	41.8	434	5 Q961H8	Q961h8 drosophila
32	41	41.8	628	5 Q9V761	Q9v761 drosophila
33	40	40.8	106	11 Q9CPS8	Q9cps8 mus musculu
34	40	40.8	120	16 Q92ER1	Q92er1 listeria in
35	40	40.8	132	16 Q926P7	Q926p7 listeria in
36	40	40.8	133	16 Q9K9F1	Q9k9f1 bacillus ha
37	40	40.8	165	5 Q9GVB4	Q9gvb4 plasmodium
38	40	40.8	239	3 Q9P8N6	Q9p8n6 cochliobolu
39	40	40.8	453	10 Q8RWR2	Q8rwr2 arabidopsis
40	40	40.8	523	10 Q94K80	Q94k80 arabidopsis
41	40	40.8	532	2 Q85154	Q85154 photorhabdu
42	40	40.8	567	5 O16108	O16108 molguila occ
43	40	40.8	568	5 Q76785	Q76785 molguila occ
44	40	40.8	596	5 Q8T6A8	Q8t6a8 caenorhabdi
45	40	40.8	598	10 Q49535	Q49535 arabidopsis

ALIGNMENTS

RESULT 1

Q9CYB9	PRELIMINARY;	PRT;	381 AA.
ID	Q9CYB9		
AC	Q9CYB9;		
DT	01-JUN-2001 (TREMREL. 17, Created)		
DT	01-JUN-2001 (TREMREL. 17, Last sequence update)		
DE	01-JUN-2002 (TREMREL. 21, Last annotation update)		
DE	Sjogren syndrome antigen B.		
GN	SSB.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=EMBRYO;		
RC	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,		
RA	Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,		
RA	Hayashizaki Y.		
RT	"Functional annotation of a full-length mouse cDNA collection."		
RL	Nature 409:685-690(2001).		
DR	EMBL; AK017822; BAB30957.1;		
DR	MGD; MGI:98423; Ssb.		

```

DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rxm; 1.
DR PRINTS; PR00302; LUPUSLA.
SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;

Query Match      84.7%; Score 83; DB 11; Length 381;
Best Local Similarity 82.4%; Pred. No. 4e-06;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALQAKICHQIQYFGQF 18
   ||| ||| ||| ||| |||
Db 12 ALEAKICHQIEYFGDF 28

RESULT 2
Q8QH15 ID Q8QH15 PRELIMINARY; PRT; 206 AA.
AC Q8QH15;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Untranslated region binding-protein.
GN UBP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA L'Ecuver T.J., Fang H.-L.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF467897; AAL76269.1;
SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFFB90B9 CRC64;

Query Match      63.38; Score 62; DB 13; Length 206;
Best Local Similarity 71.4%; Pred. No. 0.0089;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LQAKICHQIQYFG 16
   ||| ||| ||| |||
Db 13 LESKICQIEYFG 26

RESULT 3
Q9UL65 ID Q9UL65 PRELIMINARY; PRT; 671 AA.
AC Q9UL65; Q00538;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Guanine exchange factor MCG7 isoform 1 (F25B3.3 Kinase like protein).
GN CALDAG-GEFI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20493616; PubMed=10918068;
RA Clyde-Smith J., Silins G., Gartside M., Grimmond S., Etheridge M.,
RA Apolloni A., Hayward N., Hancock J.F.;
RT "Characterization of RasGRP2, a Plasma Membrane-targeted, Dual
RT Specificity Ras/Rap Exchange Factor";
RL J. Biol. Chem. 275:32260-32267(2000).
[2]
RN SEQUENCE OF 63-671 FROM N.A.
RX MEDLINE=98001089; PubMed=9341881;
RA Kedra D., Seroussi E., Fransson I., Trifunovic J., Clark M.,
RA Lagercrantz J., Blennow E., Mehlin H., Dumanski J.;
RT "The germinal center kinase gene and a novel CDC25-like gene are

```

```

RT located in the vicinity of the PYGM gene on 11q13.";
RL Hum. Genet. 100:611-619(1997).
[3]
RP SEQUENCE OF 63-671 FROM N.A.
RX MEDLINE=99007305; PubMed=9789079;
RA Kawasaki H., Springett G.M., Toki S., Canales J.J., Harlan P.,
RA Blumenstiel J.P., Chen E.J., Bany I.A., Mochizuki N., Ashbacher A.,
RA Matsuda M., Housman D.E., Graybiel A.M.;
RT "A Rap guanine nucleotide exchange factor enriched highly in the basal
RT ganglia";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13278-13283(1998).
[4]
RN SEQUENCE OF 63-671 FROM N.A.
RP Kawasaki H., Housman D.E., Graybiel A.M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
[5]
RN SEQUENCE OF 63-671 FROM N.A.
RP Kawasaki H., Housman D.E., Graybiel A.M.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[6]
RN SEQUENCE OF 63-671 FROM N.A.
RP Kawasaki H., Springett G.M., Toki S., Canales J.J., Blumenstiel J.P.,
RA Chen E.J., Bany I.A., Mochizuki N., Ashbacher A., Matsuda M.,
RA Housman D.E., Graybiel A.M.;
RT "A Novel Rap Guanine Nucleotide Exchange Factor Enriched in the Basal
RT Ganglia";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[7]
RN SEQUENCE OF 63-671 FROM N.A.
RP Silins G.U., Grimmond S., Hayward N.;
RT "Characterisation of a Novel Nucleotide Exchange Factor";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF043722; AAF07219.1;
DR EMBL; Y12336; CAA73005.1;
DR EMBL; U78170; AAD12741.1;
DR EMBL; AF081194; AAC79698.1;
DR EMBL; AF043723; AAF07220.1;
DR HSPSP; P28867; IPTQ.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000651; RasGEFN.
DR InterPro; IPR001895; RasGRF_CDC25.
DR Pfam; PF00130; DAG_PE-bind; 1.
DR Pfam; PF00036; ehand; 2.
DR Pfam; PF000617; RasGEF; 1.
DR SMART; SM00109; C1; 1.
DR SMART; SM00147; RasGEF; 1.
DR SMART; SM00229; RasGEFN; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_2.
KW Phorbol-ester binding.
SQ SEQUENCE 671 AA; 75547 MW; 67B7BD2B4F4AED4D CRC64;

Query Match      48.08; Score 47; DB 4; Length 671;
Best Local Similarity 41.28; Pred. No. 10;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 ALQAKICHQIQYFGQF 18
   ||| ||| ||| |||
Db 132 SLQVKTCHLVRYWISAF 148

RESULT 4
P72382 ID P72382 PRELIMINARY; PRT; 391 AA.
AC P72382;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cap8P.
GN CAP8P.
OS Staphylococcus aureus.

```

```

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX Staphylococcus.
RN NCBI_TaxID=1280;
RP SEQUENCE FROM N.A.
RC STRAIN-BECKER;
RX MEDLINE=96178981; PubMed=8606192;
RA Sau S., Lee C.Y.;
RT "Cloning of type 8 capsule genes and analysis of gene clusters for the
production of different capsular polysaccharides in Staphylococcus
aureus."
RL J. Bacteriol. 178:2118-2126(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BECKER;
RX MEDLINE=97197525;
RA Sau S., Sun J., Lee C.Y.;
RT "Molecular characterization and transcriptional analysis of type 8
capsule genes in Staphylococcus aureus."
RL J. Bacteriol. 179:1614-1621(1997).
DR EMBL; U73374; AAB49445.1; -
DR HSSP; P27828; 1F6D.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMS; TIGR00236; wecB; 1.
SQ SEQUENCE 391 AA; 44247 MW; 1B8D9FAA9BC76F0D CRC64;

Query Match 46.9%; Score 46; DB 2; Length 391;
Best Local Similarity 63.6%; Pred. No. 9.2;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIQYFG 16
DB 366 RICEAIEYFG 376
:| | | | |
:| | | | |

RESULT 5
ID P95709 PRELIMINARY; PRT; 391 AA.
AC P95709;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE CAP5P.
GN CAP5P.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-REYNOLDS, AND NEWMAN;
RX MEDLINE=97388587; PubMed=9445821;
RA Sau S., Bhassin N., Wann E.R., Lee J.C., Foster T.J., Lee C.Y.;
RT "The Staphylococcus aureus allelic genetic loci for serotype 5 and 8
capsule expression contain the type-specific genes flanked by common
genes."
RL Microbiology 143:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-REYNOLDS, AND NEWMAN;
RX MEDLINE=98101481; PubMed=9440531;
RA Kiser K.B., Lee J.C.;
RT "Staphylococcus aureus cap50 and cap5p genes functionally complement
mutations affecting enterobacterial common-antigen biosynthesis in
Escherichia coli."
RL J. Bacteriol. 180:403-406(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-REYNOLDS, AND NEWMAN;
RX MEDLINE=98125727; PubMed=9466251;
RA Bhassin N., Albus A., Michon F., Livolsi P.J., Park J.-S., Lee J.C.;
RT "Identification of a gene essential for O-acetylation of the

```

```

RT Staphylococcus aureus type 5 capsular polysaccharide."
RL Mol. Microbiol. 27:9-21(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-REYNOLDS, AND NEWMAN;
RA Bagga N., Wann E.R., Foster T.J., Lee J.C.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U81973; AAC46099.1; -
DR HSSP; P27828; 1F6D.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMS; TIGR00236; wecB; 1.
SQ SEQUENCE 391 AA; 44328 MW; B0105F690DB7CF1D CRC64;

Query Match 46.9%; Score 46; DB 2; Length 391;
Best Local Similarity 63.6%; Pred. No. 9.2;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIQYFG 16
DB 366 RICEAIEYFG 376
:| | | | |
:| | | | |

RESULT 6
ID Q99X57 PRELIMINARY; PRT; 391 AA.
AC Q99X57;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Capsular polysaccharide synthesis enzyme Cap5P.
GN CAPP OR SAV0164 OR SA0159.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-S. aureus (strain Mu50), and S. aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
KANAMORI M., MATSUMARU H., MARIYAMA A., MURAKAMI H., HOSIYAMA A.,
MIZUTANI-U Y., TAKAHASHI N.K., SAWANO T., INOUE R.-I., KAITO C.,
SEKIMIZU K., HIRAKAWA H., KUHARA S., GOTO S., YABUZAKI J.,
KANOHISA M., YAMASHITA A., OSHIMA K., FURUYA K., YOSHINO C., SHIBA T.,
HATTORI M., OGASAWARA N., HAYASHI H., HIRAMATSU K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003358; BAB56326.1; -
DR EMBL; AP003129; BAB41379.1; -
DR HSSP; P27828; 1F6D.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMS; TIGR00236; wecB; 1.
KW Complete proteome.
SQ SEQUENCE 391 AA; 44372 MW; D0DF5FA715BCCECC CRC64;

Query Match 46.9%; Score 46; DB 16; Length 391;
Best Local Similarity 63.6%; Pred. No. 9.2;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIQYFG 16
DB 366 RICEAIEYFG 376
:| | | | |
:| | | | |

RESULT 7
ID Q90UG9 PRELIMINARY; PRT; 608 AA.
AC Q90UG9;

```

DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Calcium and diacylglycerol-regulated guanine nucleotide exchange factor 1.
 GN RASGRP2 OR CALDAG-GEFI.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=99007305; PubMed=9789079;
 RA Kawasaki H., Springett G.M., Toki S., Canales J.J., Harlan P., Blumenstiel J.P., Chen E.J., Bany I.A., Mochizuki N., Ashbacher A., Matsuda M., Housman D.E., Graybiel A.M.;
 RA "A Rap guanine nucleotide exchange factor enriched highly in the basal ganglia.";
 RT Proc. Natl. Acad. Sci. U.S.A. 95:13278-13283(1998).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP Kawasaki H., Springett G.M., Toki S., Canales J.J., Blumenstiel J.P., Chen E.J., Bany I.A., Mochizuki N., Ashbacher A., Matsuda M., Housman D.E., Graybiel A.M.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBSJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RA Kawasaki H., Springett G.M., Toki S., Canales J.J., Blumenstiel J.P., Chen E.J., Bany I.A., Mochizuki N., Ashbacher A., Matsuda M., Housman D.E., Graybiel A.M.;
 RA "A Novel Rap Guanine Nucleotide Exchange Factor Enriched in the Basal Ganglia.";
 RT Submitted (JUL-1998) to the EMBL/GenBank/DBSJ databases.
 RL EMBL; U78171; AAD12742.1; -;
 DR EMBL; AF081193; AAC79697.1; -;
 DR HSSP; P28867; 1PTQ.
 DR MGD; MGI:1333849; Rasgrp2.
 DR InterPro; IPR002219; DAG_PE-bind.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR000651; RasGEFN.
 DR InterPro; IPR001895; RasGRF_CDC25.
 DR Pfam; PF00130; DAG_PE-bind; 1.
 DR Pfam; PF00036; efhand; 2.
 DR Pfam; PF00617; RasGEF; 1.
 DR SMART; SM00109; Cl; 1.
 DR SMART; SM00147; RasGEF; 1.
 DR SMART; SM00229; RasGEFN; 1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN 2.
 SQ SEQUENCE 608 AA; 59294 MW; DC717794CE12C2D1 CRC64;
 Query Match 46.9%; Score 46; DB 11; Length 608;
 Best Local Similarity 41.2%; Pred. No. 14;
 Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 OY 2 ALQAKICHQIQYFGF 18
 Db 70 SLQVTKCHLVRIWSAF 86
 RESULT 8
 ID O9KKL7 PRELIMINARY; PRT; 568 AA.
 AC O9KKL7;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Response regulator.
 GN VC1086.
 OS Vibrio cholerae.
 CC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RN SEQUENCE FROM N.A.

RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L., Dodson R.J., Hart D.H., Hickey E.K., Peterson J.D., Umayam L.A., Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D., Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Ragoli I., Sellers P., McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C., Fraser C.M.;
 RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.";
 RT Nature 406:477-483(2000).
 RL - SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER REGULATORY COMPONENTS OF SENSORY TRANSDUCTION SYSTEMS.
 CC EMBL; AE004434; AAF96979.1; -;
 DR HSSP; P52934; 1QMP.
 DR TIGR; VC1086;
 DR InterPro; IPR001932; PP2C-like.
 DR InterPro; IPR001789; Response_reg.
 DR Pfam; PF00072; Response_reg; 1.
 DR ProDom; PD000039; Response_reg; 1.
 DR SMART; SM00331; PP2C_SIG; 1.
 DR SMART; SM00448; REC; 1.
 KW Phosphorylation; Sensory transduction; Complete proteome.
 SQ SEQUENCE 568 AA; 64331 MW; 87A02A350D986F51 CRC64;
 Query Match 45.9%; Score 45; DB 16; Length 568;
 Best Local Similarity 42.9%; Pred. No. 19;
 Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 OY 3 LQAKICHQIQYFG 16
 Db 156 LMEENCHQVEHIFG 169
 RESULT 9
 ID O9LJ02 PRELIMINARY; PRT; 928 AA.
 AC O9LJ02;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE Similar to KIAA0731 protein.
 OS Oryza sativa (Rice).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Ehrhartoideae; Oryzae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC clone:P0499C11.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; AP001080; BAA90356.1; -;
 SQ SEQUENCE 928 AA; 103745 MW; 53FC46E24A446E4 CRC64;
 Query Match 45.9%; Score 45; DB 10; Length 928;
 Best Local Similarity 61.5%; Pred. No. 31;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 OY 3 LQAKICHQIQYFG 15
 Db 278 LRKILTQVEYVF 290
 RESULT 10
 ID O9SMD8 PRELIMINARY; PRT; 239 AA.
 AC O9SMD8;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)

DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Glycolate oxidase (EC 1.1.3.15) (Fragment).
 OS Laminaria digitata.
 OC Eukaryota; stramenopiles; Phaeophyceae; Laminariales; Laminariaceae;
 OC Laminaria.
 OX NCBI_TaxID=80365;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Moulin P., Crepinneau F., Kloareg B., Boyen C.;
 RT "Isolation and characterization of six cDNAs involved in carbohydrate
 RT biosynthesis in Laminaria digitata (Phaeophyta) - Utilization of an
 RT EST database."
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Crepinneau F., Roscoe T., Kaas R., Kloareg B., Boyen C.;
 RT "Alteration of generations in Laminaria digitata: a comparison of the
 RT gametophyte and sporophyte physiology by an EST strategy."
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ130775; CAB61335.1;
 DR HSP; P03414; IGX.
 DR InterPro; IPR003009; FMN_enzyme.
 DR InterPro; IPR000262; FMN_hydroxyac_dh.
 DR Pfam; PF01070; FMN_dh; 1
 DR PROSITE; PS00557; FMN-HYDROXY-ACID_DH; 1.
 KW Oxidoreductase.
 KW NON_TER 1
 SQ SEQUENCE 239 AA; 25573 MW; AB961336E6F20F46 CRC64;
 Query Match 44.9%; Score 44; DB 10; Length 239;
 Best Local Similarity 57.1%; Pred. No. 13; Mismatches 4; Indels 0; Gaps 0;
 Matches 8; Conservative 2;
 QY 1 AALQAKICHQIQYY 14
 II : : : : :
 Db 223 AATRAMVTHQISYY 236
 RESULT 11
 Q8RK98 PRELIMINARY; PRT; 337 AA.
 AC Q8RK98;
 DT 01-JUN-2002 (Tremblrel. 21, Created)
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
 DE Hypothetical 40.1 kDa protein.
 OS Mycoplasma hominis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2098;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=132;
 RA Boesen T.;
 RT "Gene and Protein Structure of the Mycoplasma hominis Vaa adhesin."
 RL Thesis (2001), Department of Department of Medical Microbiology and.
 DR EMBL; AJ416752; CAC95137.1;
 KW Hypothetical protein.
 SQ SEQUENCE 337 AA; 60124 MW; 6CB1C8246374CFAB CRC64;
 Query Match 43.9%; Score 43; DB 2; Length 337;
 Best Local Similarity 50.0%; Pred. No. 26; Mismatches 4; Indels 0; Gaps 0;
 Matches 6; Conservative 4;
 QY 6 KICHQIQYFGQ 17
 II : : : : :
 Db 236 KLCQIAFYCK 247
 RESULT 12
 Q97DB6 PRELIMINARY; PRT; 577 AA.
 ID Q97DB6
 AC Q97DB6;

DT 01-OCT-2001 (Tremblrel. 18, Created)
 DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Mismatch repair protein Muts-like ATPase.
 GN CAC3563.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1486;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum."
 RL J. Bacteriol. 183:4823-4838(2001).
 DR EMBL; AE007853; AAR81487.1;
 DR InterPro; IPR002106; AATRNA_ligaseII.
 DR InterPro; IPR000432; Muts_C.
 DR InterPro; IPR002863; Muts_N.
 DR Pfam; PF00488; Muts_C; 1.
 DR PRODOM; PD001263; Muts_C; 1.
 DR SMART; SM00534; MUTSd; 1.
 DR SMART; SM00533; MUTSd; 1.
 DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 577 AA; 67131 MW; DDB63C6DE85DEE7 CRC64;
 Query Match 43.4%; Score 42.5; DB 16; Length 577;
 Best Local Similarity 38.1%; Pred. No. 53; Mismatches 2; Indels 5; Gaps 1;
 Matches 8; Conservative 6;
 QY 3 LQAKICHQIQY-----YFGQF 18
 II : : : : :
 Db 53 MKAKIHKLYERFKRYWGEF 73
 RESULT 13
 O01806 PRELIMINARY; PRT; 396 AA.
 ID O01806
 AC O01806;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE C44E4.4 protein.
 GN C44E4.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium."
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Sammons L., Wohldmann P., Gillam B.;
 RT "The sequence of C. elegans cosmid C44E4."
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

RX
Citation=2005 / AACC 700332,
MEDLINE=97394467; PubMed=9252185;

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:42 ; Search time 4.95506 Seconds
(without alignments)
150.669 Million cell updates/sec

Title: US-09-836-073-3.

Perfect score: 98

Sequence: 1 AALQAKICHQIQYFGQF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	88.8	404	LA_BOVIN	P10881 bos taurus
2	87	88.8	408	LA_HUMAN	P05455 homo sapien
3	87	88.8	415	LA_RAT	P38656 rattus norv
4	83	80.7	415	LA_MOUSE	P32067 mus musculu
5	59	64.2	427	LAB_XENLA	P28049 xenopus lae
6	55	56.1	428	LAB_XENLA	P28048 xenopus lae
7	43	43.9	383	LA_AEDAL	Q28457 aedes albop
8	43	43.9	602	EX5A_BUCAL	P57530 buchnera ap
9	42	42.9	482	IFT5_HUMAN	Q13325 homo sapien
10	42	42.9	1220	DPOL_HSVB	P28858 equine herp
11	41	41.8	646	VE14_SCHPO	O13869 schizosacch
12	40	40.8	479	PRL2_ARATH	Q39190 arabidopsis
13	39.5	40.3	503	TCPT_VIBCH	P29480 vibrio chol
14	39	39.8	104	CR2_RHOVA	P00082 rhodocicrob
15	39	39.8	183	2EB2_MAIZE	P08031 zea mays (m
16	39	39.8	390	LA_DROME	P40796 drosophila
17	39	39.8	631	VE1_HPV30	Q05112 human papil
18	39	39.8	735	YD7_SCHPO	Q10432 schizosacch
19	39	39.8	926	ME19_DROME	P24087 drosophila
20	39	39.8	1221	V143_NPVAC	P24307 autographa
21	38.5	39.3	940	ADA_DROME	P09126 drosophila
22	38.5	39.3	977	A2A1_HUMAN	O95782 h adapter-r
23	38.5	39.3	977	A2A1_MOUSE	P17426 m adapter-r
24	38	38.8	251	F0L2_MOUSE	Q05685 mus musculu
25	38	38.8	305	HEM6_VIBCH	Q9KVT4 vibrio chol
26	38	38.8	786	ST5B_HUMAN	P51692 homo sapien
27	38	38.8	1088	PICO_HUMAN	Q8teq8 homo sapien
28	38	38.8	1107	MY1A_MOUSE	P46735 mus musculu
29	38	38.8	1136	MY1A_RAT	Q05096 rattus norv
30	38	38.8	1181	HAIR_RAT	P97609 rattus norv
31	38	38.8	1189	HAIR_HUMAN	Q43593 homo sapien
32	38	38.8	1596	GLI3_MOUSE	Q61602 mus musculu
33	37	37.8	227	MOAR_KLEAE	P54794 klebsiella

34	37	37.8	231	1	MTN_BRCU	O32028 bacillus su
35	37	37.8	243	1	SDGF_RAT	P24338 rattus norv
36	37	37.8	359	1	KR3_MOUSE	P51678 mus musculu
37	37	37.8	463	1	Y863_SYNY3	P73754 synchocyst
38	37	37.8	464	1	CASB_KLEOX	Q48409 klebsiella
39	37	37.8	466	1	SRO9_YEAST	P25567 saccharomyc
40	37	37.8	480	1	GLGA_RHTR	Q9cut5 rhizobium t
41	37	37.8	697	1	HRPI_ERWAM	P36654 erwini a my
42	37	37.8	715	1	ORC3_MOUSE	Q9JK30 mus musculu
43	37	37.8	719	1	FRE4_YEAST	P53746 saccharomyc
44	37	37.8	745	1	FSP1_RAT	Q63517 rattus norv
45	37	37.8	851	1	NCL1_CAEBL	P34611 caenorhabdi

ALIGNMENTS

RESULT 1
LA_BOVIN STANDARD; PRT; 404 AA.
ID LA_BOVIN
AC P10881;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
DE homolog).
GN SSB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=89202037; PubMed=2468131;
RA Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
RT sequences for RNA-binding".
RL Nucleic Acids Res. 17:2233-2244(1989).
CC -!- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
CC AND 7-2 RNAs.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
CC C-TERMINAL PART OF THE PROTEIN.
CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X13698; CAA31986.1;
DR PIR; S03849; S03849.
DR InterPro; IPR002344; Lupus.La.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KW RNA-binding; Nuclear protein; Phosphorylation.
FT DOMAIN 111 187
SQ SEQUENCE 404 AA; 46534 MW; 4EE30B5C262AD6A1 CRC64;
Query Match 88.8%; Score 87; DB 1; Length 404;

Best Local Similarity 83.3%; Pred. No. 3.1e-07;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHOIQYFGQF 18
DB 11 AALEAKICHOIEYFGDF 28

RESULT 2
LA_HUMAN
ID LA_HUMAN STANDARD; PRT; 408 AA.
AC P03455;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lupus La protein (Sjogren syndrome type B antigen (SS-B)) (La
DE ribonucleoprotein) (La autoantigen).
GN SSB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89202037; PubMed=2468131;
RA Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
RT sequences for RNA-binding.";
RL Nucleic Acids Res. 17:2233-2244(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89053970; PubMed=3192525;
RA Chambers J.C., Kenan D., Martin B.J., Keene J.D.;
RT "Genomic structure and amino acid sequence domains of the human La
RT autoantigen.";
RL J. Biol. Chem. 263:18043-18051(1988).
RN [3]
RP SEQUENCE OF 54-408 FROM N.A.
RX MEDLINE=88195081; PubMed=2452201;
RA Sturgess A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
RA Coppel R.S.;
RT "Characteristics and epitope mapping of a cloned human autoantigen
RT La.";
RL J. Immunol. 140:3212-3218(1988).
RN [4]
RP SEQUENCE OF 54-97 FROM N.A.
RX MEDLINE=85166283; PubMed=3856888;
RA Chambers J.C., Keene J.D.;
RT "Isolation and analysis of cDNA clones expressing human lupus La
RT antigen.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119(1985).
RN [5]
RP FUNCTION.
RX MEDLINE=89251617; PubMed=2470590;
RA Gottlieb E., Steitz J.A.;
RT "Function of the mammalian La protein: evidence for its action in
RT transcription termination by RNA polymerase III.";
RL EMBO J. 8:851-861(1989).
RN [6]
RP PHOSPHORYLATION.
RX MEDLINE=97207017; PubMed=9054510;
RA Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Maraie R.J.;
RT "Phosphorylation of the human La antigen on serine 366 can regulate
RT recycling of RNA polymerase III transcription complexes.";
RL Cell 88:707-715(1997).
RN [7]
RP -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
CC AND 7-2 RNAs.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- P-TM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE

C-TERMINAL PART OF THE PROTEIN.
-1- P-TM: THE N-TERMINUS IS BLOCKED.
-1- DISEASE: SERA FROM PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS
CC OPEN CONTAIN ANTIBODIES THAT REACT WITH THE NORMAL CELLULAR
CC LA PROTEIN AS IF THESE ANTIGEN WAS FOREIGN.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X13697; CAA31985.1; -
CC EMBL: J04205; AAA31885.1; -
CC PIR: A31888; A31888.
CC PIR: A22956; A22956.
CC PIR: A31273; A31273.
CC PIR: S03848; S03848.
CC PIR: S11013; S11013.
CC Genew; HGNC:11316; SSB.
CC MIM: 109090; -
CC InterPro: IPR002344; Lupus_La.
CC InterPro: IPR00504; RNA_rec_mot.
CC Pfam: PF000076; rrm; 1
CC PRINTS: PR00302; LUPUSLA.
CC SMART: SM00360; RRM; 1.
CC DR PROSITE; PS0102; RRM; 1.
CC DR PROSITE; PS00030; RRM_RNP.1; 1.
CC KW Systemic lupus erythematosus; RNA-binding; Phosphorylation;
KW Nuclear protein.
FT DOMAIN 111 187 RNA-BINDING (RRM).
FT MOD_RES 366 366 PHOSPHORYLATION (BY CK2).
SQ SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;
Query Match 88.8%; Score 87; DB 1; Length 408;
Best Local Similarity 83.3%; Pred. No. 3.2e-07;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHOIQYFGQF 18
DB 11 AALEAKICHOIEYFGDF 28

RESULT 3
LA_RAT
ID LA_RAT STANDARD; PRT; 415 AA.
AC P38656;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
DE homolog).
GN SSB OR SS-B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93246255; PubMed=7916708;
RA Semsel I., Trooster H., Bartsch H., Schwemmler M., Igloi G.L.,
RA Bachmann M.;
RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La:
RT detection of species-specific variations.";
RL Gene 126:263-268(1993).
RN -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
CC AND 7-2 RNAs.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- P-TM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE

CC AND 7-2 RNAS
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X67859; CAA48043.1; -
 CC F01494; JCI494.
 CC InterPro; IPR002344; Lupus.La.
 CC InterPro; IPR000504; RNA_rec_mot.
 CC Pfam; PF00076; rrm; 1.
 CC PRINTS; PR00302; LUPUSLA.
 CC SMART; SM00360; RRM; 1.
 CC PROSITE; PS00102; RRM; 1.
 CC RRNA-binding; Nuclear protein; Phosphorylation.
 CC FT DOMAIN 111 187 RNA-BINDING (RRM).
 CC SEQUENCE 415 AA; 47777 MW; 033FD9CC1E475F98 CRC64;
 CC
 CC Query Match 88.8%; Score 87; DB 1; Length 415;
 CC Best Local Similarity 83.3%; Pred. No. 3.2e-07;
 CC Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 1 AALQAKICHQIYFGQF 18
 CC ||:|||||:|||||
 CC DB 11 AALEAKICHQIYFGDF 28
 CC
 CC RESULT 4
 CC ID LA_MOUSE STANDARD; PRT; 415 AA.
 CC AC P32067;
 CC DT 01-OCT-1993 (Rel. 27, Created)
 CC DT 01-OCT-1993 (Rel. 27, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
 CC DE homolog).
 CC GN SSB OR SS-B.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC OX NCBI_TaxID=10090;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE-93203630; PubMed-8454877;
 CC RA Topfer F., Gordon T., McCluskey J.;
 CC RT "Characterization of the mouse autoantigen La (SS-B). Identification
 CC of conserved RNA-binding motifs, a putative ATP binding site and
 CC reactivity of recombinant protein with poly(U) and human
 CC RT autoantibodies.";
 CC RL J. Immunol. 150:3091-3100(1993).
 CC RN [2]
 CC RP SEQUENCE OF 1-11 FROM N.A.
 CC RA Groelz D., Bachmann M.;
 CC RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
 CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
 CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
 CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
 CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
 CC AND 7-2 RNAS.
 CC CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L00993; AAA39415.1; -
 CC EMBL; Y07951; CAA69249.1; -
 CC MGD; MGI:98423; Ssb.
 CC InterPro; IPR002344; Lupus.La.
 CC InterPro; IPR000504; RNA_rec_mot.
 CC Pfam; PF00076; rrm; 1.
 CC PRINTS; PR00302; LUPUSLA.
 CC SMART; SM00360; RRM; 1.
 CC PROSITE; PS00102; RRM; 1.
 CC RRNA-binding; Nuclear protein; Phosphorylation.
 CC KW RNA-binding; Nuclear protein; Phosphorylation.
 CC FT DOMAIN 111 187 RNA-BINDING (RRM).
 CC SEQUENCE 415 AA; 47756 MW; 2D75197692FDC933 CRC64;
 CC
 CC Query Match 84.7%; Score 83; DB 1; Length 415;
 CC Best Local Similarity 82.4%; Pred. No. 1.5e-06;
 CC Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 2 ALQAKICHQIYFGQF 18
 CC ||:|||||:|||||
 CC DB 12 ALEAKICHQIYFGDF 28
 CC
 CC RESULT 5
 CC ID LAB_XENLA STANDARD; PRT; 427 AA.
 CC AC P28049;
 CC DT 01-AUG-1992 (Rel. 23, Created)
 CC DT 01-AUG-1992 (Rel. 23, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen
 CC DE homolog B).
 CC GN LABI.
 CC OS Xenopus laevis (African clawed frog).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 CC OC Xenopodinae; Xenopus.
 CC OX NCBI_TaxID=8355;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE-Oocyte;
 CC RX MEDLINE-93287095; PubMed-8510143;
 CC RA Scherly D., Stutz F., Lin-Marq N., Clarkson S.G.;
 CC RT "La proteins from Xenopus laevis. cDNA cloning and developmental
 CC expression.";
 CC RL J. Mol. Biol. 231:196-204(1993).
 CC CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
 CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
 CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
 CC POLYMERASE III TRANSCRIPTS (BY SIMILARITY).
 CC CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: BARELY DETECTABLE IN STAGE I/II OOCYTES,
 CC ACCUMULATE IN STAGE III/IV OOCYTES, THEN EXHIBIT A ROUGHLY
 CC CONSTANT STEADY STATE LEVEL IN MATURE OOCYTES, EGGS, AND EARLY
 CC EMBRYOS.
 CC -1- PTM: PHOSPHORYLATED (PROBABLE).
 CC -1- MISCELLANEOUS: THERE ARE TWO FORMS OF LA, LAA AND LAB, IN XENOPUS.
 CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@lsb-slb.ch.
 CC EMBL; X68818; CAA48716.1;
 DR PIR; S28544; S28544.
 DR PIR; S33817; S33817.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 110 202 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 315 331 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 427 AA; 48995 MW; 45F3146F8934A355 CRC64;
 Query Match 60.2%; Score 59; DB 1; Length 427;
 Best Local Similarity 71.4%; Pred. No. 0.016;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Oy 3 LQAKICHOIQYFG 16
 Db 12 LDKICEIQEYFG 25
 ID LAA_XENLA STANDARD; PRT; 428 AA.
 AC P28048;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lupus La protein homolog A (La ribonucleoprotein A) (La autoantigen homolog A).
 GN LAA1.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marg N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental expression.";
 RT J. Mol. Biol. 231:196-204(1993).
 CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
 CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
 CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
 CC POLYMERASE III TRANSCRIPTS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: BARELY DETECTABLE IN STAGE I/II OOCYTES,
 CC ACCUMULATE IN STAGE III/IV OOCYTES, THEN EXHIBIT A ROUGHLY
 CC CONSTANT STEADY STATE LEVEL IN MATURE OOCYTES, EGGS, AND EARLY
 CC EMBRYOS.
 CC -1- PTM: PHOSPHORYLATED (PROBABLE).
 CC -1- MISCELLANEOUS: THERE ARE TWO FORMS OF LAA, LAA AND LAB, IN XENOPUS.
 CC -1- SIMILARITY: HIGH TO MAMMALIAN LA PROTEIN.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-slb.ch/announce/>
 CC or send an email to license@lsb-slb.ch).
 CC EMBL; X68817; CAA48715.1;
 DR PIR; S28545; S28545.

DR PIR; S33818; S33818.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 203 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 196 212 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 428 AA; 48864 MW; AEB3A38B7D2E3EC3 CRC64;
 Query Match 56.1%; Score 55; DB 1; Length 428;
 Best Local Similarity 81.8%; Pred. No. 0.077;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 6 KICHOIQYFG 16
 Db 16 KICEIQEYFG 26
 ID LAA_AEDAL STANDARD; PRT; 383 AA.
 AC Q26457;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 OS Aedes albopictus (Forest day mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
 OC Culicidae; Aedes.
 OX NCBI_TaxID=7160;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC MEDLINE=96135233; PubMed=8515178;
 RA Paradigon N., Strauss J.H.;
 RT "Mosquito homolog of the La autoantigen binds to Sindbis virus RNA.";
 RT J. Virol. 70:1173-1181(1996).
 CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTION TERMINATION BY RNA
 CC POLYMERASE III. BINDS RNA AND DNA. BINDS TO THE 3' END OF THE
 CC MINUS STRAND OF SINDBIS VIRUS RNA. THIS MAY BE SIGNIFICANT FOR
 CC SINDBIS VIRUS RNA REPLICATION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. PRIMARILY NUCLEAR, BUT SIGNIFICANT
 CC AMOUNTS ARE PRESENT IN THE CYTOPLASM.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-slb.ch/announce/>
 CC or send an email to license@lsb-slb.ch).
 CC EMBL; S80954; AAB35931.1;
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS50102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
 DR RNA-binding; Nuclear protein; DNA-binding.
 FT DOMAIN 141 228 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 228 228 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 383 AA; 44430 MW; 4E5CC6F21C40F452 CRC64;
 Query Match 43.9%; Score 43; DB 1; Length 383;
 Best Local Similarity 50.0%; Pred. No. 7.2;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LQAKICHQIYYFG 16
|:| |::|
Db 43 LEASTIROLEYFG 56

RESULT 8

EX5A_BUCAI
ID EX5A_BUCAI STANDARD; PRT; 602 AA.
AC P57530;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Exodeoxyribonuclease V alpha chain (EC 3.1.11.5).
GN RECD OR BU455.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum)
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
Buchnera sp. APS."
RL Nature 407:81-86(2000).
CC -1- FUNCTION: EXHIBITS SEVERAL CATALYTIC ACTIVITIES, INCLUDING
CC ATP-DEPENDENT EXONUCLEASE, AND ATP-STIMULATED ENDONUCLEASE,
CC ATP-DEPENDENT UNWINDING AND DNA-DEPENDENT APASE ACTIVITIES.
CC STRAND CLEAVAGE OCCURS 5' TO 3' DURING THE UNWINDING OF DUPLEX
CC DNA AT CHI SEQUENCES, WHICH LOCALLY STIMULATE RECOMBINATION
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage (in the presence of
CC ATP) in either 5'- to 3'-pr 3'- to 5'-direction to yield 5'-
CC phosphooligonucleotides.
CC -1- SUBUNIT: CONSIST OF THREE SUBUNITS; RECB, RECC AND RECD
CC (BY SIMILARITY).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

EMBL; AF001119; BAB13153.1; -

DR EMBL; AF001119; BAB13153.1; -

KW Hydrolyase; Nuclease; Exonuclease; Endonuclease; Helicase; ATP-binding;

FT NP_BIND 171 178 ATP (POTENTIAL).

FT NP_BIND 171 178 ATP (POTENTIAL).

SQ SEQUENCE 602 AA; 69494 MW; 340FFAE9BB436059 CRC64;

Query Match 43.9%; Score 43; DB 1; Length 602;

Best Local Similarity 47.1%; Pred. No. 11;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 ALQAKICHQIYYFGQF 18

|:| |::|

Db 8 AVKLIRPIDFYFSQF 24

RESULT 9

IFT5_HUMAN
ID IFT5_HUMAN STANDARD; PRT; 482 AA.
AC Q13325;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Retinoic acid- and interferon-inducible 58 kDa protein.
GN RI58.
OS Homo sapiens (Human).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

EMBL; AF001119; BAB13153.1; -

DR EMBL; AF001119; BAB13153.1; -

KW Hydrolyase; Nuclease; Exonuclease; Endonuclease; Helicase; ATP-binding;

FT NP_BIND 171 178 ATP (POTENTIAL).

FT NP_BIND 171 178 ATP (POTENTIAL).

SQ SEQUENCE 602 AA; 69494 MW; 340FFAE9BB436059 CRC64;

Query Match 43.9%; Score 43; DB 1; Length 602;

Best Local Similarity 47.1%; Pred. No. 11;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 ALQAKICHQIYYFGQF 18

|:| |::|

Db 8 AVKLIRPIDFYFSQF 24

RESULT 10

DPOL_HSVB
ID DPOL_HSVB STANDARD; PRT; 1220 AA.
AC P28858;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7).
GN 30.
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92295566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1."
RL Virology 189:304-316(1992).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
CC + [DNA](N).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98114590; PubMed=9398535;

RA Nakura T., Hirata R., Weil S.C.;

RT "A novel interferon-inducible gene expressed during myeloid

differentiation.";

RL Blood Cells Mol. Dis. 23:337-349(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Pancreas, and Spleen;

RA Strausberg R.;

RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

CC -1- INDUCTION: BY INTERFERONS.

CC -1- SIMILARITY: BELONGS TO THE IFIT FAMILY.

CC -1- SIMILARITY: CONTAINS 8 TPR REPEATS.

CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

EMBL; U34605; AAA84934.1; -

DR EMBL; U34605; AAA84934.1; -

DR EMBL; BC025786; AAH25786.1; -

DR InterPro; IPR001440; TPR.

DR Pfam; PF00515; TPR; 5.

DR SMART; SM00028; TPR; 5.

KW Repeat; TPR repeat; Interferon induction.

FT REPEAT 51 84 TPR 1.

FT REPEAT 94 127 TPR 2.

FT REPEAT 138 173 TPR 3.

FT REPEAT 181 214 TPR 4.

FT REPEAT 249 282 TPR 5.

FT REPEAT 338 371 TPR 6.

FT REPEAT 371 410 TPR 7.

FT REPEAT 435 468 TPR 8.

SQ SEQUENCE 482 AA; 55846 MW; 8045BC100384BE05 CRC64;

Query Match 42.9%; Score 42; DB 1; Length 482;

Best Local Similarity 60.0%; Pred. No. 13;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 9 HQIQYFGQF 18

||| |::|

Db 376 HQIHYHGRF 385

CC -|- SUBCELLULAR LOCATION: Nuclear.
 CC -|- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M86664; AAB02465.1; -
 CC DR PIR: D36798; DJBEC3.
 CC DR InterPro: IPR002084; DNA_pol_B.
 CC DR Pfam: PF00136; DNA_pol_B; 1.
 CC DR Pfam: PF03104; DNA_pol_B_exo; 1.
 CC DR PRINTS: PR00106; DNAPOLB.
 CC DR SMART: SM00486; POLBc; 1.
 CC DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
 CC KW Transferase; DNA-directed DNA polymerase; DNA replication;
 CC DNA-binding; Nuclear protein.
 CC SQ SEQUENCE 1220 AA; 135956 MW; 858C14DCCC71A65B CRC64;
 CC -----
 CC Query Match 42.9%; Score 42; DB 1; Length 1220;
 CC Best Local Similarity 60.0%; Pred. No. 32;
 CC Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 CC -----
 CC QY 7 ICHQIQYFG 16
 CC :||| ||||
 CC Db 248 VCHTLYYFG 257
 CC -----
 CC RESULT 11
 CC YE14_SCHPO STANDARD; PRT; 646 AA.
 CC AC O13869;
 CC DT 15-JUL-1998 (Rel. 36, Created)
 CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Putative GTP-binding protein CIB3.04C.
 CC GN SPAC1B3.04C.
 CC OS Schizosaccharomyces pombe (Fission yeast).
 CC OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC OC Schizosaccharomycetes.
 CC OX NCBI_TaxID=4896;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=972;
 CC RX MEDLINE=21848401; PubMed=11859360;
 CC RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
 CC RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 CC RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 CC RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 CC RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 CC RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 CC RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 CC RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 CC RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 CC RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 CC RA Skelton J., Simmonds R., Squares R., Squares S., Stevens K.,
 CC RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 CC RA Woodward J., Volkhardt G., Aert R., Robben J., Grynoprez B.,
 CC RA Welltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 CC RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 CC RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 CC RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,
 CC RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 CC RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 CC RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 CC RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 CC RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 CC RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -|- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC LEPA SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: Z98598; CAB11233.1; -
 CC DR HSP: P13551; IFNM.
 CC DR InterPro: IPR004161; EFTU_D2.
 CC DR InterPro: IPR00795; EF_GTPbind.
 CC DR InterPro: IPR005225; Small_GTP.
 CC DR Pfam: PF00009; GTP_EFTU; 1.
 CC DR Pfam: PF03144; GTP_EFTU_D2; 1.
 CC DR TIGRFAMs: TIGR00231; small_GTP; 1.
 CC DR PROSITE: PS00301; EFACTOR_GTP; 1.
 CC KW Hypothetical protein; GTP-binding.
 CC FT NP_BIND 65 72 GTP (POTENTIAL).
 CC FT NP_BIND 130 134 GTP (POTENTIAL).
 CC FT NP_BIND 184 187 GTP (POTENTIAL).
 CC SQ SEQUENCE 646 AA; 72683 MW; F9FA9498D384503E CRC64;
 CC -----
 CC Query Match 41.8%; Score 41; DB 1; Length 646;
 CC Best Local Similarity 46.7%; Pred. No. 25;
 CC Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 CC -----
 CC QY 3 LOAKICHQIQYFG 17
 CC :||| ||||
 CC Db 108 VRAQTCSMIYYHGQ 122
 CC -----
 CC RESULT 12
 CC PRL2_ARATH STANDARD; PRT; 479 AA.
 CC ID PRL2_ARATH
 CC AC Q39190; Q94007; Q9LUR9;
 CC DT 15-JUL-1998 (Rel. 36, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE PP1/PP2A phosphatases pleiotropic regulator PRL2.
 CC GN PRL2 OR AT3G16650 OR MGL6.11.
 CC OS Arabidopsis thaliana (Mouse-ear cress).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 CC OX NCBI_TaxID=3702;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=cv. Columbia;
 CC RX MEDLINE=98438452; PubMed=9765207;
 CC RA Nemeth K., Salcher K., Putnoky P., Bhalerao R., Koncz-Kalman Z.,
 CC RA Stankovic-Stangeland B., Bakó L., Mathur J., Okresz L., Stabel S.,
 CC RA Geigenberger P., Stitt M., Redei G.P., Schell J., Koncz C.;
 CC RT "Pleiotropic control of glucose and hormone responses by PRL1, a
 CC nuclear WD protein, in Arabidopsis";
 CC RL Genes Dev. 12:3059-3073(1998).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=cv. Columbia;
 CC RX MEDLINE=2077480; PubMed=10819329;
 CC RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
 CC RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 CC features of the regions of 4,504,864 bp covered by sixty pl and TAC
 CC clones";
 CC RL DNA Res. 7:131-135(2000).
 CC RN [3]
 CC RP SEQUENCE FROM N.A.

```

RC STRAIN-cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RL SSP consortium (Salk/Stanford/PGSC).";
CC -1- FUNCTION: PLEIOTROPIC REGULATOR OF PP1 AND PP2A PHOSPHATASES.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: BELONGS TO THE PRL1/PRL2 FAMILY OF WD-REPEAT PROTEINS.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -1- CAUTION: It is uncertain whether Met-1 or Met-3 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X82826; CAA58033.1;
DR EMBL; AB022217; BAB02756.1; ALT_SEQ.
DR EMBL; AY054181; AAL06842.1; ALT_INIT.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINRPT.
DR PRODOM: PD000018; WD40; 4.
DR SMART: SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS00682; WD_REPEATS_2; 4.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
FT REPEAT 168 198 WD 1.
FT REPEAT 210 240 WD 2.
FT REPEAT 252 282 WD 3.
FT REPEAT 293 323 WD 4.
FT REPEAT 335 364 WD 5.
FT REPEAT 377 406 WD 6.
FT REPEAT 426 456 WD 7.
SQ SEQUENCE 479 AA; 53568 MW; 301986A4AEA80670 CRC64;

Query Match 40.8%; Score 40; DB 1; Length 479;
Best Local Similarity 54.5%; Pred. No. 28;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 6 KICHIQYVFG 16
Db 48 RLCHKIQVAFG 58
:::|:|:|

RESULT 13
TCPT_VIBCH STANDARD; PRT; 503 AA.
AC P29480; Q9KTR0;
DT 01-APR-1993 (Rel. 25, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Toxin coregulated pilus biosynthesis protein T (TCP pilus
DE biosynthesis protein tcpt).
GN TCPT OR VC0835.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Classical Inaba 217561 / Serotype O1;
RX MEDLINE=93231537; PubMed=8097178;
RA Ogierman M.A., Zabihi S., Mourizios L., Manning P.A.;
RT "Genetic organization and sequence of the promoter-distal region of
RT the tcp gene cluster of Vibrio cholerae.";
RL Gene 126:51-60(1993).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emolaeve M.D., Vamathevan J., Bass S., Qin H., Raigo I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -1- FUNCTION: INVOLVED IN THE TRANSLOCATION OF THE TCPA PILIN.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE PULE/OUTE/EXEE/XPSE/XCPR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X64098; CAA45462.1;
DR EMBL; AE004168; AAF93998.1;
DR TIGR; VC0835;
DR InterPro: IPR001482; GSPIL_E.
DR Pfam: PF00437; GSPIL_E; 1.
DR PRODOM: PD000739; GSPIL_E; 1.
DR PROSITE; PS00662; T2SP_E; 1.
KW Transport; ATP-binding; Complete proteome.
FT NP_BIND 236 243 ATP (POTENTIAL).
FT CONFLICT 137 137 A -> S (IN REF. 1).
SQ SEQUENCE 503 AA; 57276 MW; ED3FD8FED579F918 CRC64;

Query Match 40.3%; Score 39.5; DB 1; Length 503;
Best Local Similarity 43.8%; Pred. No. 36;
Matches 7; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

Qy 3 LQAKICHQ-IQYVFG 17
Db 471 LKGRVCHEYFMLHFQ 486
:::|:|:|:|:|

RESULT 14
CY2_RHOVA STANDARD; PRT; 104 AA.
AC P00082;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Cytochrome C2.
OS Rhodocrobium vannielii.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Hyphomicrobium group; Rhodocrobium.
OX NCBI_TaxID=1069;
RN [1]
RP SEQUENCE.
RC STRAIN-ATCC 17100;
RX MEDLINE=76102814; PubMed=174109;
RA Ambler R.P., Meyer T.E., Kamen M.D.;
RT "Primary structure determination of two cytochromes c2: close
RT similarity to functionally unrelated mitochondrial cytochrome C.";
RL Proc. Natl. Acad. Sci. U.S.A. 73:472-475(1976).
CC -1- FUNCTION: CYTOCHROME C2 IS FOUND MAINLY IN PURPLE, NONSULFUR,
CC PHOTOSYNTHETIC BACTERIA WHERE IT FUNCTIONS AS THE ELECTRON DONOR
CC TO THE OXIDIZED BACTERIOCHLOROPHYLL IN THE PHOTOPHOSPHORYLATION
CC PATHWAY. HOWEVER, IT MAY ALSO HAVE A ROLE IN THE RESPIRATORY CHAIN
CC AND IS FOUND IN SOME NONPHOTOSYNTHETIC BACTERIA.
CC -1- PTM: BINDS ONE HEME GROUP PER MOLECULE.
DR PIR; A00074; CCRD2.

```

DR HSP; P00083; ICY.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR003088; Cyt_C1.
DR InterPro; IPR002327; Cyt_C1AB.
DR Pfam; PF00034; cytochrome_c; 1.
DR PRINTS; PR00604; CYTOCHROME_C; 1.
DR PRODOM; PD000375; Cyt_C1AB; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
KW Electron transport; Photosynthesis; Heme.
FT BINDING 14 14 HEME (COVALENT).
FT BINDING 17 17 HEME (COVALENT).
FT METAL 18 18 IRON (HEME AXIAL LIGAND).
FT METAL 80 80 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE 104 AA; 11209 MW; 1B5F2A7B657CD36F CRC64;

Query Match 39.8%; Score 39; DB 1; Length 104;
Best Local Similarity 75.0%; Pred. No. 9.9;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 QAKICHQI 11
| | | | |
Db 13 QCKICHQV 20

RESULT 15
ZEB2_MAIZE
ID ZEB2_MAIZE STANDARD; PRT; 183 AA.
AC P08031;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Zein-beta precursor (Zein 2) (16 kDa) (Zein ZC1).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87248094; PubMed=3596247;
RA Prat S., Perez-Grau L., Puigdomenech P.;
RT "Multiple variability in the sequence of a family of maize endosperm
RT proteins.";
RL Gene 52.41-49(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. W64A; TISSUE=Endosperm;
RX MEDLINE=91057131; PubMed=2243787;
RA Reina M., Guillen P., Ponte I., Boronat A., Palau J.;
RT "DNA sequence of the gene encoding the zcl protein from Zea mays W64
RT A.";
RL Nucleic Acids Res. 18:6425-6425(1990).
CC -|- FUNCTION: ZEINS ARE MAJOR SEED STORAGE PROTEINS.
CC -|- SUBCELLULAR LOCATION: ENDOSPERM PROTEIN BODIES.
CC -|- SIMILARITY: TO GLUTENIN 2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M16460; AAA35523.1; -;
CC EMBL; X53515; CAA37595.1; -;
CC PIR; B29017; B29017.
CC PIR; S12140; S12140.
CC MaizeDB; 58053; -;
CC InterPro; IPR003612; AAI.
CC InterPro; IPR001954; Glia_glutenin.
CC InterPro; IPR000480; Glutelin.
CC InterPro; IPR001768; Try/amy1_inhtr.

DR Pfam; PF00234; tryp_alpha_aml; 1.
DR PRINTS; PR00208; GLIADGLUTEN.
DR PRINTS; PR00211; GLUTELIN.
DR SMART; SM00499; AAI; 1.
KW Seed storage protein; Repeat; Multigene family; Signal.
FT SIGNAL 1 19
FT CHAIN 20 183 ZEIN-BETA.
SQ SEQUENCE 183 AA; 19558 MW; 39658BEC1151F45A CRC64;

Query Match 39.8%; Score 39; DB 1; Length 183;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ALQAKICHQIQ 12
| | | | |
Db 108 ALQOCCCHQIR 118

Search completed: April 23, 2003, 13:28:05
Job time : 6.95506 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:03 ; Search time 10.1124 Seconds
(without alignments)
171.119 Million cell updates/sec

Title: US-09-836-073-3
Perfect score: 98
Sequence: 1 AALQAKICHQIYVFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*

- 1: PIR1:*
- 2: PIR2:*
- 3: PIR3:*
- 4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	88.8	404	1 S03849	ribonucleoprotein
2	87	88.8	408	1 A31888	ribonucleoprotein
3	87	88.8	415	1 JCI494	ribonucleoprotein
4	59	60.2	427	1 S33817	ribonucleoprotein
5	55	56.1	428	1 S33818	ribonucleoprotein
6	46	46.9	391	2 H85777	capsular polysacch
7	45	45.9	568	2 C82379	response regulator
8	43	43.9	602	2 A84983	exodeoxyribonuclea
9	42.5	43.4	577	2 D97337	mismatch repair pr
10	42	42.9	396	2 T30953	hypothetical prote
11	42	42.9	482	2 G02058	retinoic acid- and
12	42	42.9	541	2 H71887	hypothetical prote
13	42	42.9	542	2 G64627	hypothetical prote
14	42	42.9	658	2 D96656	hypothetical prote
15	42	42.9	1220	1 DJBEC3	DNA-directed DNA p
16	42	42.9	1220	2 T42573	DNA-directed DNA p
17	41	41.8	259	2 T49596	related to H+-tran
18	41	41.8	336	2 S93396	HsdSB protein - M
19	41	41.8	336	2 G90566	restriction modifi
20	41	41.8	369	2 G90596	restriction-modifi
21	41	41.8	646	2 T38022	probable GTP-bindi
22	40	40.8	120	2 AF1482	hypothetical prote
23	40	40.8	133	2 A83987	ABC transporter (p
24	40	40.8	175	2 S93687	erythrocyte membra
25	40	40.8	380	2 JC5217	site-specific DNA-
26	40	40.8	598	2 T05888	hypothetical prote
27	40	40.8	658	2 T19487	hypothetical prote
28	39.5	40.3	163	2 H82733	pathogenicity-rela
29	39.5	40.3	503	2 F82275	toxin co-regulated

30	39	5	40	3	503	2	JN0524
31	39	39	39	8	104	1	CCRD2
32	39	39	39	8	166	2	T32701
33	39	39	39	8	183	2	B29017
34	39	39	39	8	346	2	A82659
35	39	39	39	8	388	2	AF0183
36	39	39	39	8	390	2	A53773
37	39	39	39	8	390	2	A53781
38	39	39	39	8	461	2	T27651
39	39	39	39	8	487	2	T27653
40	39	39	39	8	505	2	B90181
41	39	39	39	8	506	2	F85016
42	39	39	39	8	529	2	T00677
43	39	39	39	8	631	2	S36505
44	39	39	39	8	696	2	F83886
45	39	39	39	8	735	2	T41187

ALIGNMENTS

RESULT 1

S03849

ribonucleoprotein La - bovine
N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: S03849

R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.

Nucleic Acids Res. 17, 2233-2244, 1989

A:Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequ

A:Reference number: S03848; MUID:89202037; PMID:2468131

A:Accession: S03849

A:Molecule type: mRNA

A:Residues: 1-404 <CHA>

A:Cross-references: EMBL:X13698; NID:g755; PIDN:CAA31986.1; PID:g756

A>Note: part of this sequence was confirmed by protein sequencing

A:Comment: This protein associates with a variety of small RNA molecules, most of v
ay act as a transcription termination factor.

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology <RRM>

C:Keywords: blocked amino end; phosphoprotein; RNA binding

F:112-178/Domain: ribonucleoprotein repeat homology <RRM>

F:113-118/Region: RNA-binding RNP2 motif

F:151-158/Region: RNA-binding RNP1 motif

F:228-404/Domain: phosphorylated #status predicted <PHY>

Query Match 88.8%; Score 87; DB 1; Length 404;

Best Local Similarity 83.3%; Pred No. 3.6e-07;

Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIYVFGQF 18

DB 11 AALEAKICHQIEVFGDF 28

RESULT 2

A31888

ribonucleoprotein La - human

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome an

C:Species: Homo sapiens (man)

C:Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: A31888; S03848; A22956; A61051; S11013; I55553; I70205; I70206; A31273

R:Chambers, J.C.; Kenan, D.; Martin, B.J.; Keene, J.D.

J. Biol. Chem. 263, 18043-18051, 1988

A:Title: Genomic structure and amino acid sequence domains of the human La autoanti

A:Reference number: A31888; MUID:89053970; PMID:3192525

A:Accession: A31888

A:Molecule type: mRNA

A:Residues: 1-408 <CHA>

A:Cross-references: GB:J04205; NID:g178686; PIDN:AAA51885.1; PID:g178687

R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.

Nucleic Acids Res. 17, 2233-2244, 1989

A:Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequ

A:Reference number: S03848; MUID:89202037; PMID:2468131
A:Accession: S03848
A:Molecule type: mRNA
A:Residues: 1-408 <CH3>
A:Cross-references: EMBL:X13697; NID:g36414; PIDN:CAA31985.1; PID:g36415
R:Chambers, J.C.; Keene, J.D.
Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985
A:title: Isolation and analysis of cDNA clones expressing human lupus La antigen.
A:Reference number: A22956; MUID:85166283; PMID:3856888
A:Accession: A22956
A:Molecule type: mRNA
A:Residues: 45-97, 'LK' <CH3>
A:Cross-references: GB:J04205
A:Note: this sequence has been revised in reference A31888
R:Nyman, U.; Ringertz, N.R.; Pettersson, I.
Immunol. Lett. 22, 65-72, 1989
A:title: Demonstration of an amino terminal La epitope recognized by human anti-La sera
A:Reference number: A61051; MUID:89379261; PMID:2476379
A:Accession: A61051
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-19, 'E', 21-47 <NYM>
R:Sturgess, A.D.; Peterson, M.G.; McNeillage, L.J.; Whittingham, S.; Coppel, R.L.
J. Immunol. 140, 3212-3218, 1988
A:title: Characteristics and epitope mapping of a cloned human autoantigen La.
A:Reference number: S11013; MUID:88199081; PMID:2452201
A:Accession: S11013
A:Molecule type: mRNA
A:Residues: 'E', 55-287, 'V', 289-408 <STU>
A:Cross-references: EMBL:M20338; NID:g337456; PIDN:AAA36577.1; PID:g337457
R:Kohsaka, H.; Yamamoto, K.; Fujii, H.; Miura, H.; Miyasaka, N.; Nishioaka, K.; Miyamoto, J.
Clin. Invest. 85, 1566-1574, 1990
A:title: Fine epitope mapping the human SS-B/La protein: Identification of a distinct autoantigenic site
A:Reference number: I55553; MUID:90237237; PMID:1692037
A:Accession: I55553
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 81-107 <RES>
A:Cross-references: GB:M35261; NID:g338491; PIDN:AAA36652.1; PID:g338495
A:Accession: I70205
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 174-224 <RE2>
A:Cross-references: GB:M35263; NID:g338492; PIDN:AAA36653.1; PID:g338496
A:Accession: I70206
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 279-342 <RE3>
A:Cross-references: GB:M35262; NID:g338493; PIDN:AAA36654.1; PID:g338497
C:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
C:Genetics:
A:Gene: GDB:SSB
A:Cross-references: GDB:I25359; OMIM:109090
A:Map position: 2
A:Introns: 22/3; 57/2; 115/3; 185/2; 209/2; 223/3; 264/3; 380/2
C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
C:Keywords: phosphoprotein; RNA binding
F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
F:113-118/Region: RNA-binding RNP2 motif
F:151-158/Region: RNA-binding RNP1 motif
F:228-408/Domain: phosphorylated #status experimental <PHY>
Query Match 88.8%; Score 87; DB 1; Length 408;
Best Local Similarity 83.3%; Pred. No. 3.7e-07;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy 1 AALQAKICHQIQYFQG 18
    |||:|||||:|||||
Db 11 AALEAKICHQIEYFQGF 28
```

RESULT 3

JCI1494
 ribonucleoprotein La - rat
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: JCI1494; S25145
 R:Samsel, I.; Troester, H.; Bartsch, H.; Schwemmler, M.; Igloi, G.L.; Bachmann, M.
 Gene 126, 265-268, 1993
 A:Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection D
 A:Reference number: JCI1494; MUID:93246255; PMID:7916708
 A:Accession: JCI1494
 A:Molecule type: mRNA
 A:Residues: 1-415 <SEM>
 A:Cross-references: GB:X67859; NID:g55778; PIDN:CAA48043.1; PID:g55779
 A:Experimental source: liver
 C:Comment: This protein associates with a variety of small RNA molecules, most of whi
 ay act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:227-415/Domain: phosphorylated #status predicted <PHY>

Query Match 88.88; Score 87; DB 1; Length 415;
 Best Local Similarity 83.3%; Pred. No. 3.7e-07;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYVFGQF 18
 ||| ||||| ||||| |
 DB 11 AALEAKICHQIEYVFGDF 28

RESULT 4
 S33817
 ribonucleoprotein La,B - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33817; S28544
 R:Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression.
 A:Reference number: S33817; MUID:93287095; PMID:8510143
 A:Accession: S33817
 A:Molecule type: mRNA
 A:Residues: 1-427 <SCH>
 A:Cross-references: EMBL:X68818; NID:g64875; PIDN:CAA48716.1; PID:g64876
 C:Comment: This protein associates with a variety of small RNA molecules, most of whi
 ay act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:111-177/Domain: ribonucleoprotein repeat homology <RRM>
 F:112-117/Region: RNA-binding RNP2 motif
 F:150-157/Region: RNA-binding RNP1 motif
 F:227-427/Domain: phosphorylated #status predicted <PHY>

Query Match 60.2% Score 59; DB 1; Length 427;
Best Local Similarity 71.4%
Matches 10; Conservative 1; Pred. No. 0.026;
Mismatches 3; Indels 0; Gaps 0

Qy	3	LQAKICHQIQYYFG	16
Db	12	LDTKICEQIEYYFG	25

RESULTS

ribonucleoprotein La.A - African clawed frog
S33818
N;Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C;Species: *Xenopus laevis* (African clawed frog)
C;Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change
C;Accession: S33818; S28545

R.Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.

J. Mol. Biol. 231, 196-204, 1993

A:Title: La proteins from *Xenopus laevis*. cDNA cloning and developmental expression.

A:Reference number: S33817; MUID:93287095; PMID:8510143

A:Accession: S33818

A:Molecule type: mRNA

A:Residues: 1-428 <SCH>

A:Cross-references: EMBL:X68817; NID:964873; PIDN:CAA48715.1; PID:964874

A:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C:Keywords: phosphoprotein; RNA binding

F:112-178/Domain: ribonucleoprotein repeat homology <RRM>

F:113-118/Region: RNA-binding RNP2 motif

F:151-158/Region: RNA-binding RNP1 motif

F:228-428/Domain: phosphorylated #status predicted <PHY>

Query Match 56.1%; Score 55; DB 1; Length 428;

Best Local Similarity 81.8%; Pred. No. 0.13;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 KICHQIQYVFG 16

DB 16 KICEQIEYVFG 26

RESULT 6

H89777 capsular polysaccharide synthesis enzyme Cap5P [imported] - *Staphylococcus aureus* (strain

C:Species: *Staphylococcus aureus*

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: H89777

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura,

ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: H89777

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-391 <KUR>

A:Cross-references: GB:BA000018; PID:913700080; PIDN:BA841379.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: capp

C:Superfamily: lipopolysaccharide biosynthesis protein bplD

Query Match 46.9%; Score 46; DB 2; Length 391;

Best Local Similarity 53.6%; Pred. No. 4.3;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHQIQYVFG 16

DB 366 RICEAIEYVFG 376

RESULT 7

C82379 response regulator VC1A086 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)

C:Species: *Vibrio cholerae*

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: C82379

R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, J.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: C82379

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-568 <HEI>

A:Cross-references: GB:AE004434; GB:AE003853; NID:95658531; PIDN:AA696979.1; GSPDB:

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1A086

A:Map position: 2

Query Match 45.9%; Score 45; DB 2; Length 568;

Best Local Similarity 42.9%; Pred. No. 9.1;

Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 3 LOAKICHQIQYVFG 16

DB 156 LMEEMCHQVEHIFG 169

RESULT 8

A84983

exodeoxyribonuclease V (EC 3.1.11.5) 67 kD polypeptide [imported] - *Buchnera* sp. (1

C:Species: *Buchnera* sp.

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001

C:Accession: A84983

R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.

Nature 407, 81-86, 2000

A:Title: Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera*.

A:Reference number: A84930; MUID:20445173; PMID:10993077

A:Accession: A84983

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-602 <STO>

A:Cross-references: GB:AP000398; GSPDB:GN00144

A:Experimental source: strain APS

C:Genetics:

A:Gene: recD; BU455

C:Superfamily: exodeoxyribonuclease V 67K chain

C:Keywords: hydrolase

Query Match 43.9%; Score 43; DB 2; Length 602;

Best Local Similarity 47.1%; Pred. No. 21;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 ALQAKICHQIQYVFG 18

DB 8 AVKLKILRPIIDFYFSQF 24

RESULT 9

D97337 mismatch repair protein Muts-like ATPase [imported] - *Clostridium acetobutylicum*

C:Species: *Clostridium acetobutylicum*

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: D97337

R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.;

; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing *Bacteri*

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: D97337

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-577 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK81487.1; PID:915026659; GSPDB:GN00168

A:Experimental source: *Clostridium acetobutylicum* ATCC824

C:Genetics:

A:Gene: CAC3563

Query Match 43.4%; Score 42.5; DB 2; Length 577;

Best Local Similarity 38.1%; Pred. No. 25;

Matches 8; Conservative 6; Mismatches 2; Indels 5; Gaps 1;

QY 3 LOAKICHQIQY-----YFGQF 18

DB 53 MKAKIIHKLNRYERKRYWGEF 73

QY 3 LOAKICHOIQYYFGQF 18
I: : : : :
Db 520 LETRVNSLKYSAMF 535

RESULT 15

DJBEC3

DNA-directed DNA polymerase (EC 2.7.7.7) - equine herpesvirus 1 (strain Ab4p)

C:Species: equine herpesvirus 1

A:Note: host Equus caballus (domestic horse)

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 11-Jun-1999

C:Accession: D36798

R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

submitted to GenBank, March 1992

A:Description: The DNA sequence of equine herpesvirus-1.

A:Reference number: A36805

A:Accession: D36798

A:Molecule type: DNA

A:Residues: 1-1220 <TEL>

A:Cross-references: GB:M86664; NID:g330791; PIDN:AA02465.1; PID:g330822

R:Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.

Virology 189, 304-316, 1992

A:Title: The DNA sequence of equine herpesvirus-1.

A:Reference number: A41831; M01D:92295566; PMID:1318606

A:Contents: annotation; possible protein-coding frames

A:Note: neither amino acid nor nucleotide sequence is given

C:Genetics:

A:Gene: 30

C:Superfamily: herpesvirus DNA-directed DNA polymerase

C:Keywords: DNA binding; DNA replication; nucleotidyltransferase

Query Match 42.9%; Score 42; DB 1; Length 1220;
Best Local Similarity 60.0%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 ICHQIQYYFG 16
: : : : :
Db 248 VCHTTLTYFG 257

Search completed: April 23, 2003, 13:34:34
Job time : 11.1124 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:27:28 ; Search time 10.3146 Seconds
(without alignments)
139.834 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98

Sequence: 1 AALQAKICHQIYVFGQF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/FCI_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/FCIUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	98	100.0	18	9	US-09-836-073-3		Sequence 3, Appl1
2	87	88.8	18	9	US-09-836-073-1		Sequence 1, Appl1
3	87	88.8	18	9	US-09-836-073-14		Sequence 14, Appl1
4	87	88.8	460	9	US-10-102-806-695		Sequence 695, Appl1
5	83	84.7	17	9	US-09-836-073-13		Sequence 13, Appl1
6	81	82.7	18	9	US-09-836-073-9		Sequence 9, Appl1
7	79	80.6	18	9	US-09-836-073-11		Sequence 11, Appl1
8	79	80.6	18	9	US-09-836-073-12		Sequence 12, Appl1
9	78	79.6	18	9	US-09-836-073-10		Sequence 10, Appl1
10	76.5	78.1	19	9	US-09-836-073-16		Sequence 16, Appl1
11	75	76.5	18	9	US-09-836-073-2		Sequence 2, Appl1
12	75	76.5	18	9	US-09-836-073-4		Sequence 4, Appl1
13	72	73.5	18	9	US-09-836-073-7		Sequence 7, Appl1
14	71	72.4	18	9	US-09-836-073-8		Sequence 8, Appl1
15	65	66.3	18	9	US-09-836-073-15		Sequence 15, Appl1
16	63	64.3	18	9	US-09-836-073-5		Sequence 5, Appl1
17	56	57.1	18	9	US-09-836-073-6		Sequence 6, Appl1
18	50	51.0	38	9	US-09-843-676-25		Sequence 25, Appl1
19	50	51.0	38	9	US-09-766-253-25		Sequence 25, Appl1

20	50	51.0	38	9	US-09-438-486-25	Sequence 25, Appl1
21	50	51.0	38	9	US-10-053-758-25	Sequence 25, Appl1
22	50	51.0	38	9	US-10-054-295-25	Sequence 25, Appl1
23	50	51.0	38	9	US-10-054-611-25	Sequence 25, Appl1
24	49.5	50.5	37	9	US-09-843-676-24	Sequence 24, Appl1
25	49.5	50.5	37	9	US-09-766-253-24	Sequence 24, Appl1
26	49.5	50.5	37	9	US-09-438-486-24	Sequence 24, Appl1
27	49.5	50.5	37	9	US-10-053-758-24	Sequence 24, Appl1
28	49.5	50.5	37	9	US-10-054-295-24	Sequence 24, Appl1
29	49.5	50.5	37	9	US-10-054-611-24	Sequence 24, Appl1
30	45	45.9	16	9	US-09-836-073-19	Sequence 19, Appl1
31	42	42.9	186	9	US-10-013-313-6	Sequence 6, Appl1
32	41	41.8	569	10	US-09-925-300-1583	Sequence 1583, Appl1
33	40	40.8	111	10	US-09-729-835-61	Sequence 61, Appl1
34	40	40.8	2169	9	US-09-738-626-5455	Sequence 5455, Appl1
35	39	39.8	18	9	US-09-836-073-17	Sequence 17, Appl1
36	39	39.8	39	9	US-09-843-676-26	Sequence 26, Appl1
37	39	39.8	39	9	US-09-766-253-26	Sequence 26, Appl1
38	39	39.8	39	9	US-09-438-486-26	Sequence 26, Appl1
39	39	39.8	39	9	US-10-053-758-26	Sequence 26, Appl1
40	39	39.8	39	9	US-10-054-295-26	Sequence 26, Appl1
41	39	39.8	39	9	US-10-054-611-26	Sequence 26, Appl1
42	38.5	39.3	937	9	US-09-964-899-51	Sequence 51, Appl1
43	38	38.8	984	9	US-10-024-368-2	Sequence 2, Appl1
44	38	38.8	1089	9	US-10-174-590-266	Sequence 266, Appl1
45	38	38.8	1089	9	US-10-176-758-266	Sequence 266, Appl1

ALIGNMENTS

RESULT 1
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match 100.0%; Score 98; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALQAKICHQIYVFGQF 18
DB 1 AALQAKICHQIYVFGQF 18
|||||

RESULT 2
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

;; CURRENT FILING DATE: 2002-10-24
;; PRIOR APPLICATION NUMBER: 09/316,630
;; PRIOR FILING DATE: 1999-05-21
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match 88.8%; Score 87; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 6.7e-08;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
 |||||:|||||:|||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 3
US-09-836-073-14
;; Sequence 14, Application US/09836073
;; Patent No. US20020173475A1
;; GENERAL INFORMATION:
;; APPLICANT: Dasgupta, Asim
;; APPLICANT: Das, S.
;; APPLICANT: Baidya, Narayan
;; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
;; FILE REFERENCE: 22002054822
;; CURRENT APPLICATION NUMBER: US/09/836,073
;; CURRENT FILING DATE: 2002-10-24
;; PRIOR APPLICATION NUMBER: 09/316,630
;; PRIOR FILING DATE: 1999-05-21
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 14
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Bovine
US-09-836-073-14

Query Match 88.8%; Score 87; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 6.7e-08;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
 |||||:|||||:|||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 4
US-10-102-806-695
;; Sequence 695, Application US/10102806
;; Publication No. US20030054421A1
;; GENERAL INFORMATION:
;; APPLICANT: Rosen et al.
;; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;; FILE REFERENCE: P103P1C1
;; CURRENT APPLICATION NUMBER: US/10/102,806
;; CURRENT FILING DATE: 2002-03-22
;; PRIOR APPLICATION NUMBER: 09/925,298
;; PRIOR FILING DATE: 2001-08-10
;; PRIOR APPLICATION NUMBER: PCT/US00/05881
;; PRIOR FILING DATE: 2000-03-08
;; PRIOR APPLICATION NUMBER: 60/124,270
;; PRIOR FILING DATE: 1999-03-12
;; NUMBER OF SEQ ID NOS: 846
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 695
;; LENGTH: 460
;; TYPE: PRT
;; ORGANISM: Homo sapiens

US-10-102-806-695

Query Match 88.8%; Score 87; DB 9; Length 460;
Best Local Similarity 83.3%; Pred. No. 1.6e-06;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
 |||||:|||||:|||||
Db 63 AALEAKICHQIEYFGDF 80

RESULT 5
US-09-836-073-13
;; Sequence 13, Application US/09836073
;; Patent No. US20020173475A1
;; GENERAL INFORMATION:
;; APPLICANT: Dasgupta, Asim
;; APPLICANT: Das, S.
;; APPLICANT: Baidya, Narayan
;; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
;; FILE REFERENCE: 22002054822
;; CURRENT APPLICATION NUMBER: US/09/836,073
;; CURRENT FILING DATE: 2002-10-24
;; PRIOR APPLICATION NUMBER: 09/316,630
;; PRIOR FILING DATE: 1999-05-21
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13
;; LENGTH: 17
;; TYPE: PRT
;; ORGANISM: Mouse
US-09-836-073-13

Query Match 84.7%; Score 83; DB 9; Length 17;
Best Local Similarity 82.4%; Pred. No. 2.9e-07;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALQAKICHQIQYFGQF 18
 |||:|||||:|||||
Db 1 ALEAKICHQIEYFGDF 17

RESULT 6
US-09-836-073-9
;; Sequence 9, Application US/09836073
;; Patent No. US20020173475A1
;; GENERAL INFORMATION:
;; APPLICANT: Dasgupta, Asim
;; APPLICANT: Das, S. Narayan
;; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
;; FILE REFERENCE: 22002054822
;; CURRENT APPLICATION NUMBER: US/09/836,073
;; CURRENT FILING DATE: 2002-10-24
;; PRIOR APPLICATION NUMBER: 09/316,630
;; PRIOR FILING DATE: 1999-05-21
;; NUMBER OF SEQ ID NOS: 19
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 9
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match 82.7%; Score 81; DB 9; Length 18;
Best Local Similarity 87.5%; Pred. No. 6.4e-07;
Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFG 16
 |||:|||||:|||||
Db 1 AALEAKICHQIEYFG 16

```
RESULT 7
US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baluya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match      80.6%; Score 79; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 1.4e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFQGF 18
   |||:|||||:|||||
Db 1 AALEAKICHQIEYQGF 18

RESULT 8
US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baluya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match      80.6%; Score 79; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 1.4e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFQGF 18
   |||:|||||:|||||
Db 1 AALEAKICHQIEYQGF 18

RESULT 9
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baluya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match      78.1%; Score 76.5; DB 9; Length 19;
Best Local Similarity 78.9%; Pred. No. 3.7e-06;
Matches 15; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALQAKICHQIQYFQGF 18
   |||:|||||:|||||
Db 1 AALEAKICHQIEYQGF 19

RESULT 11
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baluya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2
```

Query Match 76.5%; Score 75; DB 9; Length 18;
Best Local Similarity 72.2%; Pred. No. 6.1e-06;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGF 18
DB 1 AALEAKICQIEYFGDF 18

RESULT 12
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match 76.5%; Score 75; DB 9; Length 18;
Best Local Similarity 80.0%; Pred. No. 6.1e-06;
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 QAKICHQIQYFGF 18
DB 4 EAKICHQIEYFGDF 18

RESULT 13
US-09-836-073-7
; Sequence 7, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-7

Query Match 73.5%; Score 72; DB 9; Length 18;
Best Local Similarity 81.2%; Pred. No. 1.9e-05;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFG 16
DB 1 AALEAKICHQIEYQG 16

RESULT 14
US-09-836-073-8

; Sequence 8, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-8

Query Match 72.4%; Score 71; DB 9; Length 18;
Best Local Similarity 72.2%; Pred. No. 2.7e-05;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGF 18
DB 1 AALEAKICHQIEQFGDF 18

RESULT 15
US-09-836-073-15
; Sequence 15, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Xenopus
US-09-836-073-15

Query Match 66.3%; Score 65; DB 9; Length 18;
Best Local Similarity 68.8%; Pred. No. 0.00026;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 LQAKICHQIQYFGF 18
DB 3 LDTKICEQIEYFGDF 18

Search completed: April 23, 2003, 13:38:19
Job time: 10.3146 secs

Result No.	Score		Match	Length	DB	ID	Description
	87	88					
1	87	88	8	18	4	US-09-316-630-3	Sequence 3, Appli
2	87	88	18	4	US-09-316-630-4	Sequence 4, Appli	
3	59	60	38	4	US-08-974-549A-214	Sequence 214, App	
4	50	51	38	3	US-08-851-843A-25	Sequence 25, Appl	
5	50	51	38	4	US-08-974-549A-215	Sequence 215, App	
6	50	51	38	4	US-08-854-050-25	Sequence 25, Appl	
7	50	51	38	4	US-09-430-323-25	Sequence 25, Appl	
8	49	50	37	3	US-08-851-843A-24	Sequence 24, Appl	
9	49	50	37	4	US-08-854-050-24	Sequence 24, Appl	
10	49	50	37	4	US-09-430-323-24	Sequence 24, Appl	
11	42	42	1220	2	US-08-680-326-38	Sequence 38, Appl	
12	41	41	292	4	US-09-134-001C-3515	Sequence 3515, Ap	
13	40	40	111	4	US-09-257-179-61	Sequence 61, Appl	
14	40	40	616	4	US-09-298-367B-11	Sequence 11, Appl	
15	39	39	39	3	US-08-851-843A-26	Sequence 26, Appl	
16	39	39	39	4	US-08-974-549A-216	Sequence 216, App	
17	39	39	39	4	US-08-854-050-26	Sequence 26, Appl	
18	39	39	39	4	US-09-430-323-26	Sequence 26, Appl	
19	38	39	740	4	US-09-323-872A-23	Sequence 23, Appl	
20	38	39	864	4	US-09-322-872A-28	Sequence 28, Appl	
21	38	38	305	1	US-08-420-235B-25	Sequence 25, Appl	
22	38	38	305	4	US-08-793-624-25	Sequence 25, Appl	
23	38	38	305	5	PCF-US95-1019A-25	Sequence 25, Appl	
24	38	38	984	4	US-09-287-354-2	Sequence 2, Appli	
25	38	38	1189	4	US-09-287-354-3	Sequence 3, Appli	
26	38	38	1189	4	US-09-287-354-4	Sequence 4, Appli	
27	38	38	1207	4	US-09-287-354-5	Sequence 5, Appli	

;; PRIOR FILING DATE: 1997-10-06
;; PRIOR APPLICATION NUMBER: 08/321,427
;; PRIOR FILING DATE: 1994-10-11
;; PRIOR APPLICATION NUMBER: 60/086,527
;; PRIOR FILING DATE: 1998-05-22
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4

;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
US-09-316-630-4
this peptide is biotinylated

Query Match 88.8%; Score 87; DB 4; Length 18;
Best Local Similarity 83.3%; Pred. No. 4.9e-08;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 AALQAKICHQIQYFGF 18
|||||:|||||:|||||
Db 1 AALEAKICHQIYYFGDF 18

RESULT 3

US-08-974-549A-214
; Sequence 214, Application US/08974549A
; Patent No. 6166178

GENERAL INFORMATION:

;; APPLICANT: Cech, Thomas R.
;; APPLICANT: Lingner, Joachim
;; APPLICANT: Nakamura, Toru
;; APPLICANT: Chapman, Karen B.
;; APPLICANT: Morlin, Gregg B.
;; APPLICANT: Harley, Calvin H.
;; APPLICANT: Andrews, William H.
;; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
;; NUMBER OF SEQUENCES: 727
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend and Crew LLP
;; STREET: Two Embarcadero Center, Eighth Floor
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94111-3834

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/974,549A
;; FILING DATE: 19-NOV-1997
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/844,419
;; FILING DATE: 18-APR-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/846,017
;; FILING DATE: 25-APR-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/851,843
;; FILING DATE: 06-MAY-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/854,050
;; FILING DATE: 09-MAY-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/911,312
;; FILING DATE: 14-AUG-1997

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/912,951
;; FILING DATE: 14-AUG-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/915,503
;; FILING DATE: 14-AUG-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US97/17618
;; FILING DATE: 01-OCT-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US97/17885
;; FILING DATE: 01-OCT-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph Ted
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002610US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 214:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-974-549A-214

Query Match 60.2%; Score 59; DB 4; Length 38;
Best Local Similarity 75.0%; Pred. No. 0.0039;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 7 ICHQIQYFGF 18
|||||:|||||
Db 1 ICHQXEYFGDF 12

RESULT 4

US-08-851-843A-25
; Sequence 25, Application US/08851843A
; Patent No. 6093809
;; GENERAL INFORMATION:
;; APPLICANT: Cech, Thomas R.
;; APPLICANT: Lingner, Joachim
;; APPLICANT: Nakamura, Toru
;; APPLICANT: Chapman, Karen B.
;; APPLICANT: Morlin, Gregg B.
;; APPLICANT: Harley, Calvin
;; APPLICANT: Andrews, William H.
;; TITLE OF INVENTION: No. 6093809el Telomerase
;; NUMBER OF SEQUENCES: 225
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend and Crew LLP
;; STREET: Two Embarcadero Center, 8th Floor
;; CITY: San Francisco
;; STATE: California
;; COUNTRY: United States of America
;; ZIP: 94111
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/851,843A
;; FILING DATE: 06-MAY-1997
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/846,017
;; FILING DATE: 25-APR-1997
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/844,419

```

; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-851-843A-25

```

```

Query Match          51.0%; Score 50; DB 3; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.11;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 7 ICHQIOYFG 16
  |||||
Db 1 ICEQIEYFG 10

```

```

RESULT 5
US-08-974-549A-215
; Sequence 215, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 215:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-549A-215

```

```

Query Match          51.0%; Score 50; DB 4; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.11;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 7 ICHQIOYFG 16
  |||||
Db 1 ICEQIEYFG 10

```

```

RESULT 6
US-08-854-050-25
; Sequence 25, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536

```

```

FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 35,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match 51.0%; Score 50; DB 4; Length 38;
Best Local Similarity 80.0%; Pred. No. 0.11;
Matches 8; Conservative 1; Mismatches 1; Indels

QY 7 ICHQIQYFG 16
DB 1 ICEQIEYFG 10

RESULT 8
US-08-851-843A-24
; Sequence 24, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851.843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-24

Query Match 50.5%; Score 49.5; DB 3; Length 37;
Best Local Similarity 75.0%; Pred. No. 0.13;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIYYFGQF 18
||||:|||||
Db 1 ICHQ-EYFGDF 11

RESULT 9
US-08-854-050-24
Sequence 24, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-24
Query Match 50.5%; Score 49.5; DB 4; Length 37;
Best Local Similarity 75.0%; Pred. No. 0.13;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
QY 7 ICHQIYYFGQF 18
||||:|||||
Db 1 ICHQ-EYFGDF 11
RESULT 10
US-09-430-323-24
Sequence 24, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-430-323-24

Query Match 50.5%; Score 49.5; DB 4; Length 37;
Best Local Similarity 75.0%; Pred. No. 0.13;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIQYFGQF 18
||| :|||
Db 1 ICHQ-EYFGDF 11

RESULT 11
US-08-680-326-38
Sequence 38, Application US/08680326
Patent No. 5925733
GENERAL INFORMATION:
APPLICANT: ROSE, TIMOTHY M.
APPLICANT: BOSCH, MARNIX
APPLICANT: STRAND, KURT
APPLICANT: TODARO, GEORGE J.
TITLE OF INVENTION: DNA POLYMERASE OF GAMMA HERPES VIRUSES
TITLE OF INVENTION: ASSOCIATED WITH KAPOSI'S SARCOMA AND RETROPERITONEAL
TITLE OF INVENTION: FIBROMATOSIS
NUMBER OF SEQUENCES: 152
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/680,326
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20001.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 1220 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-680-326-38

Query Match 42.9%; Score 42; DB 2; Length 1220;
Best Local Similarity 60.0%; Pred. No. 81;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 ICHQIQYFG 16
||| :|||
Db 248 VCHTLYFG 257

RESULT 12

US-09-134-001C-3515
Sequence 3515, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR FILING DATE: 1998-08-13
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-08-14
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3515
LENGTH: 292
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3515

Query Match 41.8%; Score 41; DB 4; Length 292;
Best Local Similarity 61.5%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 KICHIQIYFGQF 18
||| :|||
Db 27 KIENQIYLEQF 39

RESULT 13

US-09-257-179-61
Sequence 61, Application US/09257179
Patent No. 6410709
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 29 Human Secreted Proteins
FILE REFERENCE: P2015P1
CURRENT APPLICATION NUMBER: US/09/257,179
CURRENT FILING DATE: 1999-02-25
EARLIER APPLICATION NUMBER: PCT/US98/17709
EARLIER FILING DATE: 1998-08-27
EARLIER FILING DATE: 1998-08-27
EARLIER APPLICATION NUMBER: 60/056,270
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,271
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,247
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,073
EARLIER FILING DATE: 1997-08-29
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 61
LENGTH: 111
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (111)
OTHER INFORMATION: xaa equals stop translation
US-09-257-179-61

Query Match 40.8%; Score 40; DB 4; Length 111;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 HQIQIYFGQF 18
||| :|||
Db 86 HLPDYFGQF 95

RESULT 14

US-09-298-367B-11
; Sequence 11, Application US/09298367B
; Patent No. 6180112
; GENERAL INFORMATION:
; APPLICANT: Highlander, Sarah K.
; APPLICANT: Federova, Natalie D.
; TITLE OF INVENTION: PASTEURELLA HAEMOLYTICA VACCINE
; FILE REFERENCE: BCM-03728
; CURRENT APPLICATION NUMBER: US/09/298,367B
; CURRENT FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 08/834,455
; PRIOR FILING DATE: 1997-04-15
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Pasteurella haemolytica
US-09-298-367B-11

Query Match 40.8%; Score 40; DB 4; Length 616;
Best Local Similarity 50.0%; Pred. No. 85;
Matches 9; Conservative 3; Mismatches 4; Indels 2; Gaps 1;

QY 3 LQAK--ICHOIQYVFGF 18

Db 262 LQAKDILGHVYFLGQF 279

RESULT 15

US-08-851-843A-26
; Sequence 26, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-26

Query Match 39.8%; Score 39; DB 3; Length 39;
Best Local Similarity 60.0%; Pred. No. 7.2;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHOIQYVFG 16

Db 1 ILQVVEYFG 10

Search completed: April 23, 2003, 13:36:31
Job time : 12.2247 secs

GenCore version 5.1.4.p5.4578
Copyright. (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:23 ; Search time 28.5169 Seconds
(without alignments)
84.109 Million cell updates/sec

Title: US-09-836-073-3
Perfect score: 98
Sequence: 1 AALQAKICHQIYFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	88.8	18	AAV52200	Human la autoantigen
2	87	88.8	92	AA01351	Human secreted pro
3	87	88.8	408	AAW03716	Human autoantigen
4	87	88.8	439	ABG08417	Novel human diagno
5	87	88.8	460	AAW58987	Breast and ovarian
6	87	88.8	460	ABP41511	Human ovarian anti
7	64	65.3	21	AAW43394	La/SSB epitope 17
8	47	48.0	609	AAW87995	An alternatively s
9	47	48.0	609	AAV70961	Human Ras signall
10	47	48.0	728	AAW87994	A human MCG7 prote

11	46	46.9	608	21	AAV70960	Mouse Ras signall
12	42	42.9	175	18	AAW20649	H. pylori secreted
13	42	42.9	324	19	AAW98640	H. pylori GHPO 346
14	42	42.9	324	19	AAW71525	Helicobacter polyp
15	42	42.9	542	22	AAW46301	H. pylori HPS065 p
16	42	42.9	551	23	ABP41995	Human ovarian anti
17	42	42.9	557	23	AAW46354	H. pylori HPS065 p
18	42	42.9	2410	18	AAW19723	Cell cycle checkpo
19	42	42.9	2480	18	AAW19724	Cell cycle checkpo
20	42	42.9	2644	18	AAW13152	Human ataxia and r
21	42	42.9	2644	20	AAW84271	A human ATR protei
22	41	41.8	120	22	AAW74574	Human colon cancer
23	41	41.8	292	23	ABP38670	Staphylococcus epi
24	41	41.8	296	22	ABW60055	Drosophila melanog
25	41	41.8	544	22	AAW75090	Human colon cancer
26	41	41.8	569	21	AAW57005	Human prostate can
27	41	41.8	628	22	ABW64933	Drosophila melanog
28	41	41.8	697	22	ABW64616	Human secreted pro
29	40	40.8	111	20	AAW04314	Human secreted pro
30	40	40.8	423	21	AAW27317	Arabidopsis thalia
31	40	40.8	425	21	AAW27316	Arabidopsis thalia
32	40	40.8	477	21	AAW49578	Arabidopsis thalia
33	40	40.8	479	21	AAW49577	Arabidopsis thalia
34	40	40.8	616	19	AAW83023	Activator of leuko
35	40	40.8	2169	22	AAW91701	C glutamicum prote
36	39	39.8	40	22	ABW55660	Human testicular a
37	39	39.8	40	22	AAW94199	Human reproductive
38	39	39.8	88	22	AAW40039	Human polypeptide
39	39	39.8	88	23	ABW06216	DNA-cysteine methy
40	39	39.8	111	22	AAU19464	Human diagnostic a
41	39	39.8	111	23	ABP51350	Human MDDT SEQ ID
42	39	39.8	188	21	AAW58237	Arabidopsis thalia
43	39	39.8	188	22	AAW03032	Human polypeptide
44	39	39.8	191	22	AAW41825	Human polypeptide
45	39	39.8	232	21	AAW58236	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAV52200
ID AAV52200 standard; peptide; 18 AA.
XX
AC AAV52200;
DT 14-MAR-2000 (first entry)
XX
DE Human la autoantigen peptide (LAP).
XX

La autoantigen; LAP; internal ribosome entry site; IRES; translation;
viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
coronavirus; hepatitis virus; rhabdovirus; adenovirus; coxsackie virus;
parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
vesicular stomatitis virus.

Homo sapiens.

WO9961613-A2.

02-DEC-1999.

21-MAY-1999; 99WO-US11281.

22-MAY-1998; 98US-0086527.

(REGC) UNIV CALIFORNIA.

Das S, Dasgupta A;

WPI; 2000-062712/05.

XX

PT New yeast inhibitory peptide useful for inhibiting viral protein
 PT translation and replication -
 PS Claim 5; Page 57; 81pp; English.
 XX
 XX This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (IRNA see AA245200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rabdovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.
 XX
 XX

SQ Sequence 18 AA;

Query Match 88.8%; Score 87; DB 21; Length 18;
 Best Local Similarity 83.3%; Pred. No. 3.1e-07;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYYFGQF 18
 |||:|||||:|||||
 Db 1 AALEAKICHQIEYYFGDF 18

RESULT 2

AA01351
 ID AG01351 standard; Protein; 92 AA.

XX
 AC AAG01351;

XX 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 5432.

DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX Homo sapiens.

OS Homo sapiens.

PN EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET;

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

DR N-PSDB; AAC01357.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PT

XX Claim 13; SEQ ID 5432; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30

CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX

SQ Sequence 92 AA;

Query Match 88.8%; Score 87; DB 21; Length 92;
 Best Local Similarity 83.3%; Pred. No. 1.7e-06;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYYFGQF 18
 |||:|||||:|||||

Db 11 AALEAKICHQIEYYFGDF 28

RESULT 3

AAW03716
 ID AAW03716 standard; protein; 408 AA.

XX
 AC AAW03716;

XX 12-MAR-1997 (first entry)

XX Human autoantigen La(SS-B).

XX Autoimmune disease; La autoantigen; Sjogren's syndrome;

KW systemic lupus erythematosus; diagnosis.

XX Homo sapiens.

XX US5541291-A.

XX 30-JUL-1996.

XX 31-DEC-1984; 84US-0687908.

XX 27-MAY-1987; 87US-0054871.

XX 31-DEC-1984; 84US-0687908.

XX (UYDU-) UNIV DUKE.

XX Keene JD;

XX WPI; 1996-362015/36.

XX Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
 PT overlap syndrome - useful for diagnosis and treatment of autoimmune
 PT diseases

XX Disclosure; Columns 15-16; 21pp; English.

XX The human lupus antigen (La) is diagnostic for Sjogren's syndrome,
 CC as well as occurring in systemic lupus erythematosus patients. The
 CC La protein is clinically related to the Ro protein that is highly
 CC common among autoimmune patients. La and Ro antigens sometimes
 CC reside on the same cellular ribonucleoprotein particle; most La
 CC patients contain some Ro antibodies and vice versa. La cDNA has
 CC been isolated from a human liver library.

SQ Sequence 408 AA;

Query Match 88.8%; Score 87; DB 17; Length 408;
 Best Local Similarity 83.3%; Pred. No. 7.8e-06;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYYFGQF 18

Db 11 AALEAKICHOIYFFGDF 28

RESULT 4
ABG08417
ID ABG08417 standard; Protein; 439 AA.

AC ABG08417;
DT 13-FEB-2002 (first entry)
DE Novel human diagnostic protein #8408.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS72604.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID No 38776; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 439 AA;

Query Match 88.8%; Score 87; DB 22; Length 439;
Best Local Similarity 83.3%; Pred. No. 8.4e-06;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHOIYFFGDF 18

Db 41 AALEAKICHOIYFFGDF 58

RESULT 5

AAB58987

ID AAB58987 standard; Protein; 460 AA.

XX AAB58987;

DT 27-MAR-2001 (first entry)

DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 695.

XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neoplastic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; antitumor; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.

XX Homo sapiens.

OS WO200055173-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05881.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-611515/58.

XX N-PSDB; AAF21890.

PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases

PS Claim 11; Page 1149-11150; 1299pp; English.

CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC neoplastic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antitumor; antitumor; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiant activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.

XX Sequence 460 AA;

Query Match 88.8%; Score 87; DB 21; Length 460;
Best Local Similarity 83.3%; Pred. No. 8.8e-06;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHOIYFFGDF 18

|||||

DB 63 AALEAKICHQIEYFGDF 80

RESULT 6

ABP41511

XX ID ABP41511 standard; Protein; 460 AA.

XX AC ABP41511;

XX DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HVAF56, SEQ ID NO:2643.

XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

XX KW ovarian cancer; breast cancer; tumour; reproductive system disorder;

XX KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

XX KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;

XX KW inflammatory condition; immune disorder; blood disorder;

XX KW cardiovascular disorder; respiratory disorder; neurological disorder;

XX KW gastrointestinal disorder; urinary system disorder; drug screening;

XX KW gene therapy; chromosome mapping; forensic analysis;

XX KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;

XX KW antiinflammatory; gynaecological; reproductive.

XX OS Homo sapiens.

XX PN WO200200677-A1.

XX PD 03-JAN-2002.

XX PF 07-JUN-2001; 2001WO-US18569.

XX PR 07-JUN-2000; 2000US-209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX DR WPI; 2002-147878/19.

XX DR N-PSDB; ABQ54588.

XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,

XX PT useful in the prevention, treatment and diagnosis of cancer (e.g.

XX PT ovarian cancer), immune disorders, cardiovascular disorders and

XX PT neurological diseases -

XX PS Claim 11; SEQ ID NO 2643; 2922pp; English.

XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various ovary and/or breast-related disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present

CC sequence represents a human ovarian antigen of the invention.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 460 AA;

Query Match 88.8%; Score 87; DB 23; Length 460;

Best Local Similarity 83.3%; Pred. No. 8.8e-06;

Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIEYFGDF 18

DB 63 AALEAKICHQIEYFGDF 80

RESULT 7

AAR43394

ID AAR43394 standard; peptide; 21 AA.

XX AC AAR43394;

XX DT 12-MAY-1994 (first entry)

XX DE La/SSB epitope 17.

XX KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;

XX KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;

XX KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';

XX KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

XX OS Homo sapiens.

XX PN WO9321223-A.

XX PD 28-OCT-1993.

XX PF 13-APR-1993; 93WO-US03484.

XX PR 13-APR-1992; 92US-0867819.

XX PA (OKLA) UNIV OKLAHOMA STATE.

XX PI Harley JB;

XX DR WPI; 1993-351658/44.

XX PT New linear epitope(s) for human auto-antibodies - from the

XX PT Ro/SSA, La/SSB and Sm B/B' antigens and ribo:nucleoprotein, used

XX PT for diagnosing and treating auto-immune disorders e.g. systemic

XX PT lupus erythematosus

XX PS Claim 1; Page 30; 43pp; English.

XX CC The sequences given in AAR43391-562 are linear epitopes which are

XX CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,

XX CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'

XX CC polypeptide. These antigens are common in systemic lupus

XX CC erythematosus (SLE) and closely related disorders. The Ro/SSA family

XX CC of proteins has been shown to have several molecular forms which are

XX CC defined by the molecular weight of the antigen identified. The major

XX CC form has a molecular weight of 60 kD and two additional forms have

XX CC molecular weights of 52 and 54 kD. La/SSB is also a member of this

XX CC group of autoantibodies and binds small RNAs with a polyuridine

XX CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin

XX CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric

XX CC phosphoprotein which associates with RNA polymerase III transcripts.

XX CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and

XX CC U5 RNA. Anti-Sm antibodies may be directed against one or a

XX CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),

XX CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be

XX CC used for preventing, treating or screening autoimmune disorders,

XX CC especially SLE or Sjogrens syndrome (SS). They bind to a human

CC sequence is used for diagnosis and treatment (by protein therapy) of
CC CalDAG-GEF-associated disorders. It is also useful for developing drugs
CC and producing cell-lines or transgenic animal models for Ras-associated
CC cancers.

QY 2 ALQAKICHQIQYFGOF 18
DB 70 SLQVKTCHLVRYWISAF 86

Query Match 48.0%; Score 47; DB 21; Length 609;
Best Local Similarity 41.2%; Pred. No. 31;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

RESULT 10
AAW87994
ID AAW87994 standard; Protein; 728 AA.
AC AAW87994;
XX 15-APR-1999 (first entry)
DT A human MCG7 protein.
DE MCG4 protein; gene regulatory function; heat shock protein;
KW guanine nucleotide exchange factor protein; MCG7 protein;
KW heat shock-binding protein; MCG18 protein; zinc finger protein;
KW cancer.
XX Homo sapiens.
OS
XX WO9853061-A1.
PN 26-NOV-1998.
XX 22-MAY-1998; 98WO-AU00380.
PF 22-JAN-1998; 98AU-0001460.
PR 23-MAY-1997; 97AU-0006972.
PR 23-MAY-1997; 97AU-0006973.
PR 23-MAY-1997; 97AU-0006974.
PR 22-JAN-1998; 98AU-0001458.
PR 22-JAN-1998; 98AU-0001459.
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
PA Gartside M, Grimmond S, Hancock J, Hayward N, Silins G;
PI WPI; 1999-070146/06.
DR N-PSDB; AAX04552.
XX New gene-expression regulatory genes, mcg4, mcg7, and mcg18 - encode
PT a zinc finger protein, a GEF, and a heat shock or heat shock binding
PT protein, useful to detect and treat cancer
XX Claim 5; Fig 13a; 80pp; English.
PS
XX The present sequence represents a MCG7 protein. The protein has gene
CC regulatory functions, and has homology to a heat shock protein or
CC heat shock-binding protein. The specification also describes MCG4,
CC which is homologous to guanine nucleotide exchange factor protein,
CC and MCG18, which is homologous to a zinc finger protein.
CC Detection of mutations in the MCG genes can be used to identify the
CC propensity for various types of cancer, and to treat, arrest, or
CC otherwise ameliorate, the effects of a cancer in an animal or bird.
XX
SQ Sequence 728 AA;

Query Match 48.0%; Score 47; DB 20; Length 728;
Best Local Similarity 41.2%; Pred. No. 37;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 ALQAKICHQIQYFGOF 18
DB 189 SLQVKTCHLVRYWISAF 205

RESULT 11
AAW70960
ID AAW70960 standard; Protein; 608 AA.
XX AAW70960;
XX 09-AUG-2000 (first entry)
DT Mouse Ras signalling pathway associated protein CalDAG-GEFI.
DE Mouse; Ras signalling pathway; CalDAG-GEFI; calcium; diacylglycerol;
KW guanine nucleotide exchange factor; Rap1A; diagnosis; treatment;
KW CalDAG-GEF-associated disorder; drug; transgenic animal model;
KW Ras-associated cancer; protein therapy.
XX Mus musculus.
OS
XX WO200024768-A2.
PN 04-MAY-2000.
XX 22-OCT-1999; 99WO-US24826.
PF 23-OCT-1998; 98US-0105507.
PR 16-NOV-1998; 98US-0108685.
XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
PA Kawasaki H, Graybiel A, Housman D;
PI WPI; 2000-350690/30.
DR N-PSDB; AAD00310.
XX Isolated nucleic acid comprises nucleotide sequence encoding protein
PT selected from normal or mutant CalDAG-guanine nucleotide exchange
PT factor
XX Claim 64; Page 86-88; 128pp; English.
PS The present sequence is a mouse CalDAG-guanine nucleotide
XX exchange factor1 (GEFI). It has substrate specificity for
CC Rap1A, dual binding domains for calcium and diacylglycerol (DAG), and is
CC highly expressed in brain, particularly in brain basal ganglia pathways
CC and their axon-terminal regions. Expression of CalDAG-GEFI activates
CC Rap1A and inhibits Ras-dependent activation of the extracellular-signal
CC regulated kinase/mitogen-activated protein (ERK/MAP) kinase cascade in
CC 293T cells. The CalDAG-GEF proteins play an important role in determining
CC the relative activation of Ras and Rap1 signalling induced by calcium and
CC DAG mobilisation in brain and haematopoietic organs. The present
CC sequence is used for diagnosis and treatment (protein therapy) of
CC CalDAG-GEF-associated disorders. It is also useful for developing drugs
CC and producing cell-lines or transgenic animal models for Ras-associated
CC cancers.
XX
SQ Sequence 608 AA;

Query Match 46.9%; Score 46; DB 21; Length 608;
Best Local Similarity 41.2%; Pred. No. 45;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 ALQAKICHQIQYFGOF 18
DB 70 SLQVKTCHLVRYWISAF 86

RESULT 12
AAW20649

ID AAW20649 standard; protein; 175 AA.
 AC AAW20649;
 XX
 DT 14-JUL-1997 (first entry)
 XX
 DE H. pylori secreted or periplasmic protein, O2cp20821orf12.
 XX
 KW Cytoplasmic; vaccine; prevention; treatment; infection; identification;
 binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
 KW duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
 XX
 OS Helicobacter pylori.
 XX
 PN W09640893-AL.
 XX
 PD 19-DEC-1996.
 XX
 PF 06-JUN-1996; 96WO-US09122.
 XX
 PR 01-APR-1996; 96US-0630405.
 PR 07-JUN-1995; 95US-0487032.
 XX
 PA (ASTR) ASTRA AB.
 XX
 PI Berglindh OT, Smith D, Mellgaard BL;
 DR WPI; 1997-052306/05.
 DR N-PSDB; AAT67902.
 XX
 PT Helicobacter pylori nucleic acid sequences and related
 PT polypeptide(s) - useful for vaccines to treat or prevent H. pylori
 PT infection, and to detect Helicobacter
 XX
 PS Claim 72; Page 1072; 1481pp; English.
 XX
 CC The present sequence is a H. pylori secreted or periplasmic protein.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds,
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The genomic sequence of H. pylori (ATCC 55679) was determined from
 CC overlapping contigs generated by mechanically shearing the bacterial
 CC DNA. The sequences were analysed for ORF of at least 180 nucleotides,
 CC and the predicted coding regions defined by computer evaluation. To
 CC identify likely H. pylori antigens for vaccine development, the amino
 CC acid sequences predicted from various ORF were analysed for significant
 CC homology to other known or exported membrane proteins. Having identified
 CC and determined the sequences of interest, particular regions can be
 CC isolated from H. pylori by PCR amplification for recombinant polypeptide
 CC production, e.g. in E. coli hosts.
 XX
 SQ Sequence 175 AA;
 Query Match 42.9%; Score 42; DB 18; Length 175;
 Best Local Similarity 50.0%; Pred. No. 54;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AALQAKICHQIQ 12
 Db 134 AEVDKVKCHQVE 145
 RESULT 13
 AAW98640
 ID AAW98640 standard; Protein; 324 AA.
 XX
 AC AAW98640;
 XX
 DT 31-MAR-1999 (first entry)
 XX
 DE H. pylori GHPO 346 protein.
 XX
 KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;

KW peptic ulcer disease.
 XX
 OS Helicobacter pylori.
 XX
 PN W09843478-AL.
 XX
 PD 08-OCT-1998.
 XX
 PF 01-APR-1998; 98WO-US06371.
 XX
 PR 29-JUL-1997; 97US-0902615.
 PR 01-APR-1997; 97US-0833457.
 PR 24-JUN-1997; 97US-0881227.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
 XX
 PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
 XX
 DR WPI; 1998-542293/46.
 DR N-PSDB; AAX14359.
 XX
 PT New isolated Helicobacter polynucleotides - used to develop products
 PT for the diagnosis, prevention and treatment of Helicobacter
 PT infections and gastrointestinal diseases
 XX
 PS Claim 8; Page 1309-1310; 2054pp; English.
 XX
 CC This sequence represents a Helicobacter pylori GHPO protein of the
 CC invention. The polypeptides can be used for preventing or treating
 CC Helicobacter infections, and gastroduodenal diseases associated with
 CC these infections, including acute, chronic, and atrophic gastritis, and
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
 CC used for the production of antibodies. The products can also be used for
 CC detection and diagnosis.
 XX
 SQ Sequence 324 AA;
 Query Match 42.9%; Score 42; DB 19; Length 324;
 Best Local Similarity 50.0%; Pred. No. 1e+02;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AALQAKICHQIQ 12
 Db 123 AEVDKVKCHQVE 134
 RESULT 14
 AAW71525
 ID AAW71525 standard; Protein; 324 AA.
 XX
 AC AAW71525;
 XX
 DT 09-NOV-1998 (first entry)
 XX
 DE Helicobacter polypeptide GHPO 346.
 XX
 KW GHPO 346; infection; therapy; diagnosis; vaccine; gastritis;
 KW ulcer.
 XX
 OS Helicobacter pylori.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 315 /note- "encoded by NTC"
 FT Misc-difference 316 /note- "encoded by NTC"
 XX
 PN W09821225-AL.
 XX
 PD 22-MAY-1998.
 XX
 PS 14-NOV-1997; 97WO-US21353.

```

XX PR 29-JUL-1997; 97US-0902615.
XX PR 14-NOV-1996; 96US-0749051.
XX PR 01-APR-1997; 97US-0831309.
XX PR 01-APR-1997; 97US-0833457.
XX PR 01-APR-1997; 97US-0834705.
XX PR 24-JUN-1997; 97US-0881227.
XX PR
XX (HUMA-) HUMAN GENOME SCI INC.
XX PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX PA (INNR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX XX
XX PI Al-Garawi A, Haas R, Kleanthous H, Meyer T, Miller C;
XX PI Odenbreit S, Tomb J;
XX PR
XX WPI; 1998-297855/26.
XX DR N-PSDB; AAV52060.
XX PR
XX Helicobacter polynucleotide and polypeptide sequences - useful to
XX PT treat or prevent gastrointestinal infection
XX XX
XX PS Claim 1; Page 220-221; 362pp; English.
XX PR
XX This claimed Helicobacter pylori polypeptide, designated GHP0 346,
XX CC can be used in vaccination methods for preventing or treating
XX CC Helicobacter infection. 85 Helicobacter polypeptides (see
XX CC AAW71474-W71558) are claimed, as well as isolated polynucleotides
XX CC (see AAV52009-93) that encode them. The invention also provides:
XX CC methods for producing these Helicobacter polypeptides in
XX CC recombinant host systems, and related expression cassettes, vectors
XX CC and transformed or transfected host cells; live vaccine vectors
XX CC that contain the polynucleotides of the invention and which can be
XX CC used to prevent or treat Helicobacter infection; therapeutic and/or
XX CC prophylactic methods involving administration of polynucleotide
XX CC molecules, polypeptides or monospecific antibodies; methods for
XX CC detecting the presence of Helicobacter in samples using e.g.
XX CC the polypeptides or monospecific antibodies; and methods for
XX CC purifying the polypeptides by antibody-based affinity
XX CC chromatography.
XX PR
XX SQ Sequence 324 AA;
XX
XX Query Match 42.9%; Score 42; DB 19; Length 324;
XX Best Local Similarity 50.0%; Pred. No. 1e+02;
XX Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 AALQAKICHQIQ 12
XX Db 123 AEVQDKVCHQVE 134
XX | : | : | : | :
XX
XX RESULT 15
XX AAB46301
XX ID AAB46301 standard; Protein; 542 AA.
XX AC AAB46301;
XX XX
XX DT 05-APR-2001 (first entry)
XX XX
XX DE H. pylori HPS065 protein.
XX XX
XX KW Microbial infection; antibacterial; Helicobacter pylori infection;
XX KW vaccine; screening.
XX XX
XX OS Helicobacter pylori.
XX XX
XX PN WO200073502-A2.
XX XX
XX PD 07-DEC-2000.
XX XX
XX PF 31-MAY-2000; 2000WO-EP05024.
XX XX
XX PR 31-MAY-1999; 99DE-1024965.

```

```

PR 17-JUN-1999; 99DE-1027740.
PR 21-JUL-1999; 99DE-1034029.
XX XX
XX PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX PA (CREA-) CREATOGEN GMBH.
XX XX
XX PI Apfel H, Fuchs TM, Gibbs CP, Hueck CJ, Meyer TF;
XX PR
XX WPI; 2001-0499948/06.
XX DR N-PSDB; AAF25578.
XX PR
XX Preparing an agent for diagnosis or control of microbial infection,
XX PT useful particularly against Helicobacter, based on identification of
XX PT essential genes in defective mutants -
XX PR
XX PS Claim 37; Page 201-203; 366pp; German.
XX PR
XX This invention describes a novel preparation of an agent (A) for
XX CC detection, prevention and/or treatment of microbial infection by:
XX CC (i) identifying essential genes (I) and corresponding polypeptides
XX CC (II); (ii) identifying compounds that are directed against (II) and
XX CC inactivate the microbe; (iii) testing these for suitability for use; and
XX CC (iv) formulating selected (A). Identifying essential genes (I) comprises
XX CC preparation of gene-deficient microorganisms by conditional antisense
XX CC inhibition (CAI) and/or subtractive recombination mutagenesis (SRM),
XX CC then determining viability and/or survival of the deficient organisms.
XX CC The products of the invention have antibacterial activity. (A) (which may
XX CC be a nucleic acid (Ia), vector or host cell containing (Ia), derived
XX CC polypeptide (Iia), or fragments, (Iia)-specific antibodies or their
XX CC fragments or an inhibitor of (Iia)) are particularly used for diagnosis,
XX CC treatment or prevention of infection by Helicobacter pylori. Particularly
XX CC (Ia) and (Iia) are used in DNA, subunit or live vaccines. The method
XX CC identifies essential genes, including those that have homologs in other
XX CC species, so identified (A) should have a broad spectrum of activity. Many
XX CC gene-deficient cells can be screened quickly, in an automated process,
XX CC and the identified genes can be used for screening without purification.
XX PR
XX SQ Sequence 542 AA;
XX
XX Query Match 42.9%; Score 42; DB 22; Length 542;
XX Best Local Similarity 50.0%; Pred. No. 1.7e+02;
XX Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 AALQAKICHQIQ 12
XX Db 123 AEVQDKVCHQVE 134
XX | : | : | : | :
XX
XX Search completed: April 23, 2003, 13:27:10
XX Job time : 30.5169 secs

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:34:48 ; Search time 20.0225 Seconds
(Without alignments)
185,234 Million cell updates/sec

Title: US-09-836-073-3
Perfect score: 98
Sequence: 1 AALQAKICHQIQYFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 9297

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp-archaea.*
2: sp-bacteria.*
3: sp-fungi.*
4: sp-human.*
5: sp-invertebrate.*
6: sp-mammal.*
7: sp-mhc.*
8: sp-organelle.*
9: sp-phage.*
10: sp-plant.*
11: sp-rodent.*
12: sp-virus.*
13: sp-vertebrate.*
14: sp-unclassified.*
15: sp-rvirus.*
16: sp-bacteriaph.*
17: sp-archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	31.6	24	5 Q9BM09	Q9bm09 spongilla 1
2	28.5	29.1	23	5 Q95S95	Q95s95 drosophila
3	28	28.6	24	2 Q05616	Q05616 staphylococ
4	27	27.6	13	5 Q917F8	Q917f8 drosophila
5	27	27.6	24	6 Q28430	Q28430 gorilla gor
6	27	27.6	24	6 Q28781	Q28781 pongo pygma
7	25	25.5	16	2 Q9R963	Q9r963 helicobacte
8	25	25.5	18	4 Q16028	Q16028 homo sapien
9	25	25.5	19	6 Q9BDD6	Q9bdd6 ovis aries
10	25	25.5	23	13 Q9PRV2	Q9prv2 gallus gall
11	25	25.5	25	13 Q9PRV1	Q9prv1 gallus gall
12	24.5	25.0	24	9 Q38270	Q38270 bacterioph
13	24	24.5	12	10 Q9M433	Q9m433 lotus japon
14	24	24.5	17	11 P97758	P97758 mus musculu
15	24	24.5	19	2 Q9R4B9	Q9r4b9 streptococ
16	24	24.5	19	8 Q9XMB5	Q9xmb5 aegilops ta

17	24	24.5	19	12	083273	Q83273 cucumber mo
18	24	24.5	24	7	Q9TNS7	Q9tns7 homo sapien
19	24	24.5	24	7	Q9TNS6	Q9tns6 homo sapien
20	24	24.5	25	5	Q9TNC6	Q9tnc6 drosophila
21	24	24.5	25	8	Q9TGB8	Q9tgb8 alnus crisp
22	24	24.5	25	8	Q9TGB7	Q9tgb7 alnus glut
23	24	24.5	25	8	Q9TGB6	Q9tgb6 alnus marit
24	24	24.5	25	8	Q9TGB5	Q9tgb5 betula alle
25	24	24.5	25	8	Q9TGB4	Q9tgb4 betula glan
26	24	24.5	25	8	Q9TGB3	Q9tgb3 betula papy
27	24	24.5	25	8	Q9TGB2	Q9tgb2 betula verr
28	24	24.5	25	8	Q9TGB1	Q9tgb1 betula pube
29	24	24.5	25	8	Q9TGB0	Q9tgb0 corylus ave
30	24	24.5	25	8	Q9TGA9	Q9tga9 corylus col
31	24	24.5	25	8	Q9TGA8	Q9tga8 corylus cor
32	24	24.5	25	8	Q9TGA7	Q9tga7 ostrya virg
33	24	24.5	25	8	Q9TGA6	Q9tga6 quercus rub
34	24	24.5	25	8	Q9TGB3	Q9tgb3 carpinus ca
35	24	24.5	25	13	Q90Z49	Q90z49 haplochromi
36	23	23.5	8	4	Q9BY5	Q9by5 homo sapien
37	23	23.5	8	6	Q9BFA0	Q9bfa0 macaca mula
38	23	23.5	8	6	Q9BF99	Q9bf99 hylobates c
39	23	23.5	9	12	Q9E1U7	Q9e1u7 hepatitis b
40	23	23.5	13	12	Q9E1V4	Q9e1v4 hepatitis b
41	23	23.5	13	12	Q9E1V3	Q9e1v3 hepatitis b
42	23	23.5	13	12	Q9E1V2	Q9e1v2 hepatitis b
43	23	23.5	13	12	Q9E1V1	Q9e1v1 hepatitis b
44	23	23.5	13	12	Q9E1V0	Q9e1v0 hepatitis b
45	23	23.5	13	12	Q9E1U9	Q9e1u9 hepatitis b

ALIGNMENTS

RESULT 1					
Q9BM09	PRELIMINARY;	PRT;	24	AA.	
ID	Q9BM09				
AC	Q9BM09;				
DT	01-JUN-2001 (Tremblrel. 17, Created)				
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)				
DT	01-JUN-2001 (Tremblrel. 17, Last annotation update)				
DE	Gypsy-like reverse transcriptase (Fragment).				
OS	Spongilla lacustris (Freshwater sponge).				
OC	Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;				
OC	Haplosclerida; Spongillidae; Spongilla.				
OX	NCBI_TaxID=6055;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TRANSPOSON-GRT-G7 RETROTRANSPOSON;				
RX	MEDLINE-20570504; PubMed-11121049;				
RA	Arkhipova I., Meselson M.;				
RT	"Transposable elements in sexual and asexual taxa.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).				
DR	EMBL; AY013997; AAG59969.1; -				
KW	RNA-directed DNA polymerase.				
FT	NON_TER 1				
FT	NON_TER 24				
SQ	SEQUENCE 24 AA; 2754 MW; B588512EDE7A251A CRC64;				

Query Match 31.6%; Score 31; DB 5; Length 24;
Best Local Similarity 30.8%; Pred. No. 2.3e+02;
Matches 4; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY	3	LQAKICHQIQYF	15
Db	11	LMDTVCHGLEFVF	23

RESULT 2					
Q95S95	PRELIMINARY;	PRT;	23	AA.	
ID	Q95S95				
AC	Q95S95;				
DT	01-DEC-2001 (Tremblrel. 19, Created)				

DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE GM05003P.

GN ANTP OR BG:DS07700.1 OR CG1028.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,
 RL SUBMIT (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY060900; AA128448.1;
 DR FlyBase; FBgn0000095; Antp.
 SQ SEQUENCE 23 AA; 2820 MW; 6F16D52A47F69139 CRC64;

Query Match 29.18; Score 28.5; DB 5; Length 23;

Best Local Similarity 53.8%; Pred. No. 6e+02;

Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 3 LOAKI-CHOIOY 14

Db 11 IOIKHPHIVY 23

RESULT 3

ID Q05616 PRELIMINARY; PRT; 24 AA.

AC Q05616;

DT 01-NOV-1996 (TRENBLrel. 01, Created)

DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)

DT 01-NOV-1996 (TRENBLrel. 08, Last annotation update)

DE 3-dehydroquinase synthase (EC 4.6.1.3) (Fragment).

GN AROB

OS Staphylococcus aureus.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Staphylococcus.

OX NCBI_TaxID=1280;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=8325-4;

RX MEDLINE=93381456; PubMed=8371108;

RA O'Connell C.M., Pattee P., Foster T.J.;

RT "Sequence and mapping of the aroA gene of Staphylococcus aureus 8325-4.";

RL J. Gen. Microbiol. 139:1449-1460(1993).

CC -1- CATALYTIC ACTIVITY: 7-PHOSPHO-3-DEOXY-ARABINO-HEPTULOSONATE = 3-

CC DEHYDROQUINATE + ORTHOPHOSPHATE.

CC -1- PATHWAY: SECOND STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE

CC AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).

CC -1- SUBUNIT: MONOMER (BY SIMILARITY).

DR EMBL: L05004; AA71896.1;

KW Aromatic amino acid biosynthesis; Lyase.

FT NON_TER 1

SQ SEQUENCE 24 AA; 2857 MW; 1C4C8D20F7A4856E CRC64;

Query Match

Best Local Similarity 50.0%; Score 28; DB 2; Length 24;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 CHQIOYF 15

Db 16 CEOLKTYF 23

RESULT 4

Q917F8

ID Q917F8 PRELIMINARY; PRT; 13 AA.

AC Q917F8;

DT 01-MAR-2001 (TRENBLrel. 16, Created)

DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)

DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)

DE CG18278 protein (Fragment).

GN CG18278.

OS Drosophila melanogaster (fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

MDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA Ananides P.G., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos G.L.G.,

RA Arif J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,

RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J., Yao Q.A.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

DR EMBL: AE003818; AAG2274.1;

DR FlyBase; FBgn0033836; CG18278.

FT NON_TER 1

SQ SEQUENCE 13 AA; 1554 MW; 5D52855A93735EB7 CRC64;

Query Match 27.6%; Score 27; DB 5; Length 13;

Best Local Similarity 36.4%; Pred. No. 6.3e+02;

Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 LOAKICHOIQ 13

Db 2 LOAGLCITNF 12

RESULT 5

Q28430

ID Q28430 PRELIMINARY; PRT; 24 AA.

AC Q28430;

DT 01-NOV-1996 (TRENBLrel. 01, Created)

DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Relaxin (Fragment).
 GN RLX.
 OS Gorilla gorilla (gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 OX NCBI_TaxID=9593;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RG YK;
 RA Evans B.B.A., Fu P., Tregear G.G.W.;
 RT "Characterisation of primate relaxin genes.";
 RL J. Mol. Endocrinol. 0:0-0(1993).
 CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 DR EMBL; Z27228; CAAB1742.1; -;
 DR HSSP; P04090; 6RLX.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 24 AA; 2673 MW; 5615D4B92C1DFD24 CRC64;

Query Match 27.6%; Score 27; DB 6; Length 24;
 Best Local Similarity 55.6%; Pred. No. 1.1e+03;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AALQAKICH 9
 :||| |
 Db 4 SALANKCCH 12

RESULT 6
 Q28781 PRELIMINARY; PRT; 24 AA.
 AC Q28781;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Relaxin (Fragment).
 GN RLX.
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RG YK 287;
 RA Evans B.B.A., Fu P., Tregear G.G.W.;
 RT "Characterisation of primate relaxin genes.";
 RL J. Mol. Endocrinol. 0:0-0(1993).
 CC -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
 DR EMBL; Z27229; CAAB1743.1; -;
 DR HSSP; P04090; 6RLX.
 DR InterPro; IPR004825; Ins/IGF/relax.
 DR Pfam; PF00049; Insulin; 1.
 DR PROSITE; PS00262; INSULIN; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 24 AA; 2673 MW; 5615D4B92C1DFD24 CRC64;

Query Match 27.6%; Score 27; DB 6; Length 24;
 Best Local Similarity 55.6%; Pred. No. 1.1e+03;
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AALQAKICH 9
 :||| |
 Db 4 SALANKCCH 12

RESULT 7
 Q9R963 PRELIMINARY; PRT; 16 AA.
 ID Q9R963

AC Q9R963;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE PepC (Fragment).
 GN FEPC.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FE31;
 RX MEDLINE=98453456; PubMed=9780260;
 RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,
 RA Kuriyama M.;
 RT "Full-length sequence analysis of the vacA gene from cytotoxic and
 RT noncytotoxic Helicobacter pylori.";
 RL J. Infect. Dis. 178:1391-1398(1998).
 DR EMBL; AF049623; AAD04263.1; -;
 FT NON_TER 1
 FT NON_TER 1
 FT NON_TER 16
 SQ SEQUENCE 16 AA; 1938 MW; 3C33D403AEE61428 CRC64;

Query Match 25.5%; Score 25; DB 2; Length 16;
 Best Local Similarity 44.4%; Pred. No. 1.7e+03;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 QIQYFGQF 18
 :||| |
 Db 7 QVEYYAFNF 15

RESULT 8
 Q16028 PRELIMINARY; PRT; 18 AA.
 AC Q16028;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE OCRL-1 protein (Fragment).
 GN OCRL-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93278398; PubMed=8504307;
 RA Leachey A.M., Charnas L.R., Nussbaum R.L.;
 RT "Nonsense mutations in the OCRL-1 gene in patients with the
 RT oculocerebrorenal syndrome of Lowe.";
 RL Hum. Mol. Genet. 2:461-463(1993).
 DR EMBL; S62084; AAD13933.1; -;
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 18 AA; 2055 MW; 61356C825F7AF33D CRC64;

Query Match 25.5%; Score 25; DB 4; Length 18;
 Best Local Similarity 36.4%; Pred. No. 1.9e+03;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 AALQAKICHQI 11
 :||| |
 Db 1 SAYDPRICRQL 11

RESULT 9
 Q9BDD6 PRELIMINARY; PRT; 19 AA.
 ID Q9BDD6
 AC Q9BDD6;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE Choline/ethanolamine kinase isolog (Fragment).
GN CHOLINE KINASE.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Price N.T., Cameron J.M., van der Leij F.R., Zammit V.A.;
RT "Analysis of the promoter region for the ovine carnitine
RT palmitoyltransferase IB gene";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ288906; CAC36323.1; -
DR EMBL; AJ288907; CAC36324.1; -
KW Kinase.
FT NON_TER 1
SQ SEQUENCE 19 AA; 2328 MW; 938A7B7F1636318C CRC64;

Query Match 25.5%; Score 25; DB 6; Length 19;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIQYFQG 17
Db :||| |
1 RFQYFQG 8

RESULT 10
Q9PRV2 PRELIMINARY; PRT; 23 AA.
ID Q9PRV2
AC Q9PRV2;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Elastase 1 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE
RX MEDLINE=95066525; PubMed=7976036;
RA Tuscik F., Polanowski A., Guyonnet V., Long P.L., Travis J.;
RT "Affinity purification of chicken pancreas proteinases and their N-
RT terminal amino-acid sequences.";
RL Acta Biochim. Pol. 41:174-177(1994).
DR HSP; P00772; LEA1.
SQ SEQUENCE 23 AA; 2491 MW; 441147DDB26A987C CRC64;

Query Match 25.5%; Score 25; DB 13; Length 23;
Best Local Similarity 56.7%; Pred. No. 2.4e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 IQYFQG 16
Db :||| |
18 LOYFQG 23

RESULT 11
Q9PRV1 PRELIMINARY; PRT; 25 AA.
ID Q9PRV1
AC Q9PRV1;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Elastase 2 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.

OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE
RX MEDLINE=95066525; PubMed=7976036;
RA Tuscik F., Polanowski A., Guyonnet V., Long P.L., Travis J.;
RT "Affinity purification of chicken pancreas proteinases and their N-
RT terminal amino-acid sequences.";
RL Acta Biochim. Pol. 41:174-177(1994).
DR HSP; P00772; LEA1.
SQ SEQUENCE 25 AA; 2635 MW; BB72141147DDB26A CRC64;

Query Match 25.5%; Score 25; DB 13; Length 25;
Best Local Similarity 66.7%; Pred. No. 2.6e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 IQYFQG 16
Db :||| |
18 LOYFQG 23

RESULT 12
Q38270 PRELIMINARY; PRT; 24 AA.
ID Q38270
AC Q38270;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Bacteriophage lambda mutant truncated regulatory protein (PyrI) gene
DE from E.coli, 3' end (Fragment).
OS Bacteriophage lambda.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10710;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86143826; PubMed=3912513;
RA Cunin R., Jacobs A., Charlier D., Crabeel M., Herve G., Glansdorff N.,
RA Pierard A.;
RT "Structure-function relationship in allosteric aspartate
RT carbamoyltransferase from Escherichia coli: I. Primary structure of a
RT pyrI gene encoding a modified regulatory subunit.";
RL J. Mol. Biol. 186:707-713(1985).
DR EMBL; M28579; AAA32252.1; -
DR HSP; P00478; AATC.
FT NON_TER 1
SQ SEQUENCE 24 AA; 2905 MW; FD9349DF6F8159D0 CRC64;

Query Match 25.0%; Score 24.5; DB 9; Length 24;
Best Local Similarity 38.5%; Pred. No. 3e+03;
Matches 5; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 2 ALQAKICHQIY 14
Db :||| |
8 ALKCYKER-EFY 19

RESULT 13
Q9M433 PRELIMINARY; PRT; 12 AA.
ID Q9M433
AC Q9M433;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE ENOD40-1 protein.
GN ENOD40-1.
OS Lotus japonicus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20427395; PubMed=10975655;

RA Fletakis E., Kavroulakis N., Quaedvlieg N.E.M., Spalink H.P.,
 RA Dimou M., Roussis A., Katinakis P.,
 RT "Lotus japonicus contains two distinct ENOD40 genes that are expressed
 in symbiotic, non-symbiotic and embryonic tissues.",
 RL Mol. Plant Microbe Interact. 13:987-994(2000).
 DR EMBL; AJ271787; CAB92978.1;
 SQ SEQUENCE 12 AA; 1403 MW; 3C6955187CB046C3 CRC64;

Query Match 24.5%; Score 24; DB 10; Length 12;
 Best Local Similarity 44.4%; Pred. No. 1.9e+03;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 6 KICHQIQY 14
 | : | : | :
 Db 2 KICWQISIH 10

RESULT 14

P97758 PRELIMINARY; PRT; 17 AA.
 AC P97758;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Dbm28 protein (Fragment).
 GN H-2D.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J-H-2BM28;
 RA MEDLINE=97188477; PubMed=9037062;
 RX Yun T.J., Melvold R.W., Pease L.R.;
 RT "A complex major histocompatibility complex D locus variant generated
 by an unusual recombination mechanism in mice."
 RL Proc. Natl. Acad. Sci. U.S.A. 94:1384-1389(1997).
 DR EMBL; U83304; AAB41021.1;
 DR InterPro; IPR001039; MHC_I.
 DR Pfam; PF00129; MHC_I; 1.
 FT NON_TER 1
 FT NON_TER 17
 SQ SEQUENCE 17 AA; 1956 MW; F3149F377C16F196 CRC64;

Query Match 24.5%; Score 24; DB 11; Length 17;
 Best Local Similarity 37.5%; Pred. No. 2.7e+03;
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 9 HQIQYIFG 16
 | : | : | :
 Db 3 HTLQWYMG 10

RESULT 15

Q9R4B9 PRELIMINARY; PRT; 19 AA.
 AC Q9R4B9;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE Surface trypsin-resistant LADDERING protein (Fragment).
 OS Streptococcus sp.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID=1306;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=97047708; PubMed=8926097;
 RA Lachenaue C.S., Madoff L.C.;
 RT "A protective surface protein from type V group B streptococci shares
 N-terminal sequence homology with the alpha C protein.";
 RL Infect. Immun. 64:4255-4260(1996).

SQ SEQUENCE 19 AA; 2008 MW; F44CC89ED774EA3C CRC64;
 Query Match 24.5%; Score 24; DB 2; Length 19;
 Best Local Similarity 50.0%; Pred. No. 3e+03;
 Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 AALQAKICHQIQ 12
 | : | : | :
 Db 7 ATLNTKITRNQI 18

Search completed: April 23, 2003, 13:47:14
 Job time : 21.0225 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:33:13 ; Search time 4.65169 Seconds
(without alignments)
160.495 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98
Sequence: 1 AALQAKICHQIYFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1520

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	24.5	11	1 CXL1_CONMR	P58807 conus marmo
2	24	24.5	12	1 TIN2_HOFTI	P82652 hoplobatrach
3	24	24.5	13	1 CXL1_CONMR	P58810 conus marmo
4	24	24.5	18	1 MLE_SCYCA	P01206 scyllorhinu
5	24	24.5	23	1 XYL1_ACIGB	P46365 acinetobact
6	24	24.5	25	1 CXOB_CONMA	P05485 conus magus
7	24	24.5	25	1 SMP_RAT	P80968 rattus norv
8	23	23.5	16	1 MLE_SQUAC	P01207 squalus aca
9	23	23.5	20	1 TLI8_SPIOL	P82536 spinacia ol
10	23	23.5	25	1 ANDT_ANDAU	P56684 androctonus
11	22	22.4	8	1 ALI7_CARMA	P81820 carcinus ma
12	22	22.4	9	1 ALI1_CARMA	P81820 carcinus ma
13	22	22.4	14	1 LECB_FSOCC	P22584 psophocarpu
14	22	22.4	17	1 TPIS_FINPS	P81666 pinus pinas
15	22	22.4	21	1 BTX_ATRBI	P80163 attractaspis
16	22	22.4	21	1 SRUD_ATRBN	P13211 attractaspis
17	21	21.4	18	1 OBP_LYMDI	P34173 lymantria d
18	21	21.4	19	1 HBB2_URORH	P18992 uronastyr h
19	21	21.4	23	1 PRO3_DACGL	P18690 dactylis gl
20	20	20.4	14	1 CXAL_CONCN	P56973 conus conso
21	20	20.4	15	1 PC20_BRANA	P81096 brassica na
22	20	20.4	16	1 CXA2_CONMA	P56636 conus magus
23	20	20.4	20	1 COG4_CHIOP	P34156 chironocete
24	20	20.4	21	1 NDK_CANAL	Q9ur66 candida alb
25	20	20.4	22	1 LPI_TRIWA	P24335 trimeresuru
26	20	20.4	22	1 LP2_TRIWA	P58930 trimeresuru
27	20	20.4	23	1 NUO5_SOLTU	P80262 solanum tub
28	20	20.4	25	1 NEU1_RIG	P34964 sus scrofa
29	19	19.4	10	1 GKN2_CHEPR	P80878 chelyosoma
30	19	19.4	10	1 TRNB_ONCMY	P28500 oncorhynch
31	19	19.4	11	1 TIN4_HOFTI	P82654 hoplobatrach
32	19	19.4	11	1 TKNA_CHICK	P19850 gallus gall
33	19	19.4	11	1 TKNA_HORSE	P01290 equus cabal

34	19	19.4	12	1 NO40_SESRO	O24369 sesbania ro
35	19	19.4	12	1 TIN3_HOFTI	P82653 hoplobatrach
36	19	19.4	14	1 ADF_TENMO	P82965 tenebrio mo
37	19	19.4	14	1 MAST_VESLE	P01514 vespula lew
38	19	19.4	15	1 IRBP_CRISP	P12665 cricetidae
39	19	19.4	15	1 PGT5_PELAC	P80564 pelobacter
40	19	19.4	18	1 SPAH_HELAN	P81098 helianthus
41	19	19.4	20	1 LPYR_BACST	P05858 bacillus st
42	19	19.4	20	1 PNW2_PHONI	Q9twr5 phoneutria
43	19	19.4	21	1 TRNC_CARAU	P25421 carassius a
44	19	19.4	22	1 ATP6_COTJA	P50681 coturnix co
45	19	19.4	22	1 VGLG_RABVA	P15199 rabies viru

ALIGNMENTS

RESULT 1				
ID	CXL1_CONMR	STANDARD	PRT	11 AA
AC	P58807			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lambda-conotoxin CMrVIA.			
OS	Conus marmoreus (Marble cone).			
OC	Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;			
OC	Neogastropoda; Conoidea; Conidae; Conus.			
OX	NCBI_TaxID=42752;			
RN	[1]			
RP	SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.			
RC	TISSUE=Venom;			
RX	MEDLINE=20564325; PubMed=10988292;			
RA	Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,			
RA	Seow K.T., Bay B.-H.;			
RT	"Lambda-conotoxins, a new family of conotoxins with unique disulfide			
RT	pattern and protein folding. Isolation and characterization from the			
RT	venom of Conus marmoreus.";			
RL	J. Biol. Chem. 275:39516-39522(2000).			
CC	- FUNCTION: Inhibits the neuronal noradrenaline transporter.			
CC	- SUBCELLULAR LOCATION: Secreted.			
CC	- TISSUE SPECIFICITY: Expressed by the venom duct.			
CC	- MASS SPECTROMETRY: MW=1237.93; MWERR=0.21; METHOD=Electrospray.			
CC	- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.			
KW	Neurotoxin; Toxin; Hydroxylation.			
FT	DISULFID 2 11			
FT	DISULFID 3 8			
FT	MOD_RES 10 10			
SQ	SEQUENCE 11 AA; 1226 MW; 277AAC60B7232B58 CRC64;			
Query Match 24.5%; Score 24; DB 1; Length 11;				
Best Local Similarity 75.0%; Pred. No. 3.9e+02;				
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;				
QY	6 KICH 9			
DB	6 KICH 9			
RESULT 2				
ID	TIN2_HOFTI	STANDARD	PRT	12 AA
AC	P82652			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Tigerinus-2.			
OS	Hoplobatrachus tigerinus (Indian bull frog) (Rana tigerina).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Neobatrachia; Ranidae;			
OC	Hoplobatrachus.			
OX	NCBI_TaxID=103373;			
RN	[1]			

```

MLB_SCVCA                                PRT:      18 AA.
ID      MLB_SCVCA      STANDARD;
AC      P01206;
DT      21-JUL-1986 (Rel. 01, Created)
DT      21-JUL-1986 (Rel. 01, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Melanotropin beta (Beta-MSH).
OS      Scylliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC      Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC      Scyllorhinidae; Scyllorhinus.
OX      NCBI_TaxID=7830;
RN      [1]
RP      SEQUENCE.
RX      MEDLINE=75113445; PubMed=4452470;
RT      "Love R.M., Pickering B.T.;
RT      "A beta-MSH in the pituitary gland of the spotted dogfish
RT      (Scylliorhinus canicula): isolation and structure.";
RL      Gen. Comp. Endocrinol. 24:398-404(1974).
RC      -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
RD      PIR; A01470; MTFDEC.
RK      Hormone.
KW      SEQUENCE      18 AA;      2195 MW;      353C93A30C314DF4 CRC64;
SQ
Query Match      24.5%;      Score 24;      DB 1;      Length 18;
Best Local Similarity      44.4%;      Pred. No. 6.2e+02;
Matches      4;      Conservative      1;      Mismatches      4;      Indels      0;      Gaps
QY      10 QIQYFGQF 18
      :| | | |
DB      2 ZIBYKMGHF 10

RESULT 5
XYCL_ACIGB                                STANDARD;                                PRT:      23 AA.
ID      XYCL_ACIGB
AC      P46365;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Benzaldehyde dehydrogenase [NAD+] I (EC 1.2.1.28) (Fragment).
OS      Acinetobacter genosomp. 11.
OC      Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC      Acinetobacter.
OX      NCBI_TaxID=106649;
RN      [1]
RP      SEQUENCE.
RX      STRAIN=ATCC 11171 / NCIB 8250 / CIP 63.46 / B94;
RC      MEDLINE=91113163; PubMed=1989592;
RT      Chalmers R.M., Keen J.N., Fewson C.A.;
RT      "Comparison of benzyl alcohol dehydrogenases and benzaldehyde
RT      dehydrogenases from the benzyl alcohol dehydrogenases and benzaldehyde
RT      Acinetobacter calcoaceticus and from the TOL-plasmid-encoded toluene
RT      pathway in Pseudomonas putida. N-terminal amino acid sequences, amino
RT      acid compositions and immunological cross-reactions.";
RL      Biochem. J. 273:99-107(1991).
RC      -1- CATALYTIC ACTIVITY: Benzaldehyde + NAD(+) + H(2)O = benzoate +
CC      NADH.
CC      -1- SUBUNIT: HOMOTETRAMER.
CC      -1- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
CC      InterPro: IPR002086; Aldehyde_dehydr.
DR      PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; PARTIAL.
DR      PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; PARTIAL.
KW      Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.
FT      NON_TER      23
FT      SEQUENCE      23 AA;      2750 MW;      50496D4930DF9E40 CRC64;

Query Match      24.5%;      Score 24;      DB 1;      Length 23;
Best Local Similarity      55.6%;      Pred. No. 7.8e+02;
Matches      5;      Conservative      1;      Mismatches      3;      Indels      0;      Gaps
QY      3 LOAKICHQI 11
      :| | | |

```

DB 3 IOTKIEQI 11

RESULT 6

ID CXOB_CONMA STANDARD; PRT; 25 AA.
AC P05485;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DE Omega-conotoxin MVIIB (SNX-159).
OS Conus magus (Magus cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6492;
RN [1]
RP SEQUENCE.
RX MEDLINE=87299637; PubMed=2441741;
RA Olivera B.M., Cruz L.J., de Santos V., Lecheminant G.W., Griffin D.,
RA Zeikus R.D., McIntosh J.M., Galyean R., Varga J., Gray W.R.,
RA Rivier J.E.;
RT "Neuronal calcium channel antagonists. Discrimination between calcium
channel subtypes using omega-conotoxin from Conus magus venom.";
RL Biochemistry 26:2086-2090(1987).
CC - FUNCTION: Omega-conotoxins act at presynaptic membranes, they bind
CC and block voltage-sensitive calcium channels (VSCC).
CC - SUBCELLULAR LOCATION: Secreted.
CC - TISSUE SPECIFICITY: Expressed by the venom duct.
CC - SIMILARITY: BELONGS TO THE O-SUPERFAMILY OF CONOTOXINS. OMEGA-TYPE
CC FAMILY.

DR PIR; B34115; B34115.

DR HSP; JH0701; JH0701.

DR HSP; P05484; IMV1.

KW Presynaptic neurotoxin; Neurotoxin; Toxin; Calcium channel inhibitor;

KW Amidation.

FT DISULFID 1 16

FT DISULFID 8 20

FT DISULFID 15 25

FT MOD_RES 25 25 AMIDATION.

SQ SEQUENCE 25 AA; 2626 MW; E4B9CE5EFAA3734D CRC64;

Query Match 24.5%; Score 24; DB 1; Length 25;

Best Local Similarity 50.0%; Pred. No. 8.5e+02;

Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 8 CHQIQY 13

II: I

DB 8 CHRTSY 13

RESULT 7

ID SMBP_RAT STANDARD; PRT; 25 AA.
AC P80968;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE SM-11044 binding protein (Fragments).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RX STRAIN=Histar;
RX Sugawara T., Matsuzaki-Fujita M., Guillaume J.-L., Camoin L.,
RA Morooka S., Strosberg A.D.;
RT "Characterization of a novel iodoctanopindolol and SM-11044 binding
RT protein, which may mediate relaxation of depolarized rat colon
RT tonus.";
RL J. Biol. Chem. 272:21244-21252(1997).
CC - FUNCTION: MAY MEDIATE RELAXATION OF DEPOLARIZED COLON TONUS.

CC - IT BINDS IODOCTANOPINDOLOL AND SM-11044.

CC - MISCELLANEOUS: THE ORDER OF THE FRAGMENTS IS UNKNOWN.

FT NON_TER 1 1

FT UNSURE 6 6 OR Y.

FT NON_CONS 18 19

FT NON_TER 25 25

SQ SEQUENCE 25 AA; 3177 MW; D14F0CB9B778C2CB CRC64;

Query Match 24.5%; Score 24; DB 1; Length 25;

Best Local Similarity 57.1%; Pred. No. 8.5e+02;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 12 QYVGFQF 18

II: I

DB 15 QYFYPXF 21

RESULT 8

ID MLB_SQUAC STANDARD; PRT; 16 AA.
AC P01207;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE Melanotropin beta.
DE Melanotropin beta.
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalia; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE.
RX MEDLINE=75127390; PubMed=4375978;
RA Bennett H.P.J., Lowry P.J., McMartin C., Scott A.P.;
RT "Structural studies of alpha-melanocyte-stimulating hormone and a
RT novel beta-melanocyte-stimulating hormone from the neurointermediate
RT lobe of the pituitary of the dogfish Squalus acanthias.";
RL Biochem. J. 141:439-444(1974).
CC - SIMILARITY: BELONGS TO THE POMC FAMILY.
DR PIR; A01471; WTDPBS.
KW Hormone.

SQ SEQUENCE 16 AA; 1939 MW; 993AF43C8A640A0E CRC64;

Query Match 23.5%; Score 23; DB 1; Length 16;

Best Local Similarity 66.7%; Pred. No. 8.2e+02;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 13 YVFGQF 18

II: I

DB 5 YKFGHF 10

RESULT 9

ID TL18_SPIOL STANDARD; PRT; 20 AA.
AC P82536;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Thylakoid lumenal 18 kDa protein (P18) (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllaceae; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RX Kieselbach T., Bystedt M., Schroeder W.P.;
RL Submitted (MAY-2000) to the SWISS-PROT data bank.
CC - SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen.
KW Chloroplast; Thylakoid.
FT NON_TER 20 20

SQ SEQUENCE 20 AA; 2192 MW; 752C21963F49FA64 CRC64;

Query Match 23.5%; Score 23; DB 1; Length 20;
 Best Local Similarity 27.3%; Pred. No. 1e+03;
 Matches 3; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 3 QAKICHQIQY 13
 |||: : : :
 Db 7 LQSKVTNKVVF 17

RESULT 10
 ANDT_ANDAU STANDARD; PRT; 25 AA.
 ID ANDT_ANDAU STANDARD; PRT; 25 AA.
 AC P56684; P81616;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Androctonus.
 OS Androctonus australis hector (Sahara scorpion).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Euthoidea; Euthoidea; Androctonus.
 OX NCBI_TaxID=70175;
 RN [1]
 RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
 RC TISSUE-Hemolymph;
 RX MEDLINE=97094646; PubMed=8939880;
 RA Ehret-Sabatier L., Loew D., Goyffon M., Fehlbaum P., Hoffmann J.A.,
 van Dorsseleer A., Bulet P.;
 RT "Characterization of novel cysteine-rich antimicrobial peptides from
 scorpion blood.";
 RL J. Biol. Chem. 271:29537-29544(1996).
 RN [2]
 RP SYNTHESIS OF D-AMINO ACID ENANTIOMER, AND CHARACTERIZATION.
 RX MEDLINE=20115101; PubMed=10642525;
 RA Hetru C., Letellier L., Oren Z., Hoffmann J.A., Shai Y.;
 RT "Androctonin, a hydrophilic disulphide-bridged non-haemolytic
 anti-microbial peptide: a plausible mode of action.";
 RL Biochem. J. 345:653-664(2000).
 RN [3]
 RP STRUCTURE BY NMR.
 RX MEDLINE=20025109; PubMed=10563585;
 RA Mandard N., Sy D., Maufrais C., Bonmatin J.M., Bulet P., Hetru C.,
 Vovelle F.;
 RT "Androctonin, a novel antimicrobial peptide from scorpion Androctonus
 australis: solution structure and molecular dynamics simulations in
 the presence of a lipid monolayer.";
 RL J. Biomol. Struct. Dyn. 17:367-380(1999).
 CC -1- FUNCTION: ACTIVE AGAINST BOTH BACTERIA (GRAM-POSITIVE AND GRAM-
 NEGATIVE) AND FILAMENTOUS FUNGI. ACTS ON THE MEMBRANE OF THE
 BACTERIAL CELLS. IT DESTABILIZE A MEMBRANE BY MODIFYING ITS
 PROPERTIES
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MASS SPECTROMETRY: MW=3076.; METHOD=Electrospray.
 DR PDB; IC26; 12-JAN-00.
 KW Antibiotic; Fungicide; 3D-structure.
 FT DISULFID 4 20
 FT DISULFID 10 16
 SQ SEQUENCE 25 AA; 3081 MW; D4183D6EA5AA50AA CRC64;

Query Match 23.5%; Score 23; DB 1; Length 25;
 Best Local Similarity 57.1%; Pred. No. 1.2e+03;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 QAKICHQ 10
 | : : : :
 Db 6 QIKICRR 12

RESULT 11
 AL17_CARMA STANDARD; PRT; 8 AA.
 ID AL17_CARMA STANDARD; PRT; 8 AA.
 AC P81820;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinus maenas (Common shore crab) (Green crab).
 OS Carcinus maenas (Common shore crab); Mandibulata; Pancrustacea; Crustacea;
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
 OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
 OC Brachyura; Eubrachyura; Portunoidae; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 8 8
 FT AMIDATION (POTENTIAL).
 SQ SEQUENCE 8 AA; 858 MW; C82879D5AB46D865 CRC64;

Query Match 22.4%; Score 22; DB 1; Length 8;
 Best Local Similarity 80.0%; Pred. No. 1.1e+05;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 QYFPG 16
 ||| :
 Db 3 QYFSG 7

RESULT 12
 AL17_CARMA STANDARD; PRT; 9 AA.
 ID AL17_CARMA STANDARD; PRT; 9 AA.
 AC P81814;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinus maenas (Common shore crab) (Green crab).
 OS Carcinus maenas (Common shore crab); Mandibulata; Pancrustacea; Crustacea;
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
 OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
 OC Brachyura; Eubrachyura; Portunoidae; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 9 9
 FT AMIDATION.
 SQ SEQUENCE 9 AA; 927 MW; 832D79CDBC46D861 CRC64;

Query Match 22.4%; Score 22; DB 1; Length 9;
 Best Local Similarity 80.0%; Pred. No. 1.1e+05;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 QYFPG 16
 ||| :
 Db 4 QYAFG 8

RESULT 13
 LECB_PSOSC STANDARD; PRT; 14 AA.
 ID LECB_PSOSC STANDARD; PRT; 14 AA.
 AC P22584;
 DT 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Basic lectin B1 (Fragment).
OS Psophocarpus scandens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
OC Psophocarpus.
OX NCBI_TaxID=3890;
RN [1]
RN SEQUENCE.
RC TISSUE-Seed;
RA Kott A.A.;
RT Isolation and characterization of the lectins from the seeds of
RT Psophocarpus scandens.
RL Phytochemistry 27:2847-2855(1988).
CC -1- SUBUNIT: DIMER OF TWO IDENTICAL OR NEARLY IDENTICAL SUBUNITS OF
CC ABOUT 32000 APPARENT MW.
CC -1- PTM: CONTAINS ABOUT 9% CARBOHYDRATE. HETEROGENEITY IN CARBOHYDRATE
CC COMPOSITION MAY RESULT IN DIFFERENT ISOLECTIN FORMS.
CC -1- SIMILARITY: WITH P.TETRAGONALOBUS BASIC LECTINS IN N-TERMINAL
CC SEQUENCE, AA COMPOSITION, ERYTHROCYTE AND SUGAR SPECIFICITIES.
DR PIR; PA0007; PA0007.
KW Lectin; Glycoprotein.
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1732 MW; D804CE43B487C549 CRC64;

Query Match 22.4%; Score 22; DB 1; Length 14;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 11 IQYFQGF 18
DB 3 ISFNQGF 10

RESULT 14
ID TPIS_PINPS STANDARD; PRT; 17 AA.
AC P81666;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Triosephosphate isomerase, cytosolic (EC 5.3.1.1) (TIM) (Fragments).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RN SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT Separation and characterization of needle and xylem maritime pine
RT proteins.
RL Electrophoresis 20:1098-1108(1999).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glycerone
CC phosphate.
CC -1- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- INDUCTION: BY WATER STRESS.
CC -1- MISCELLANEOUS: IN PLANTS, THERE ARE TWO TYPES OF TPIS, CYTOSOLIC
CC AND PLASTID.
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN
CC (SPOT N139) IS: 5.9, ITS MW IS: 24 kDa.
CC -1- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
DR InterPro: IPR000652; Triophos_Isomr.
DR PROSITE: PS00171; TIM; PARTIAL.
KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
KW Pentose shunt.
FT NON_TER 1 1

FT NON_CONS 9 10
FT NON_TER 17 17
SQ SEQUENCE 17 AA; 2053 MW; 183DB41757AF13CB CRC64;

Query Match 22.4%; Score 22; DB 1; Length 17;
Best Local Similarity 20.0%; Pred. No. 1.3e+03;
Matches 2; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 7 ICHQIQYFVG 16
DB 4 VCVEQLFFVG 13

RESULT 15
ID BTX_ATRBI STANDARD; PRT; 21 AA.
AC P80163;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bibratoxin (BTX).
OS Atractaspis bibroni (Southern bibron's) (Stiletto snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Atractaspididae; Atractaspis.
OX NCBI_TaxID=8601;
RN [1]
RN SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=93106214; PubMed=8416802;
RA Becker A., Dowdle E.B., Hechler U., Kauser K., Donner P.,
RA Schleuning W.-D.;
RT Bibratoxin, a novel member of the endothelin/sarafotoxin peptide
RT family, from the venom of the burrowing asp Atractaspis bibroni.
RL FEBS Lett. 315:100-103(1993).
CC -1- FUNCTION: VASOCONSTRICTOR ACTIVITY. THESE TOXINS CAUSE CARDIAC
CC ARREST PROBABLY AS A RESULT OF CORONARY VASOSPASM.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE ENDOTHELIN/SARAFOTOXIN FAMILY.
DR PIR; S27039; S27039.
DR HSSP; P13208; 1SRB.
DR InterPro: IPR001928; Endothlin_tox.
DR InterPro: IPR003642; Sara/bib_toxin.
DR Pfam; PF00322; endothelin; 1.
DR PRINTS; PR00365; ENDOTHELIN.
DR ProDom; PD004740; Sara/bib_toxin; 1.
DR SMART; SM00272; END; 1.
DR PROSITE; PS00270; ENDOTHELIN; 1.
KW Vasoconstrictor; Toxin.
FT DISULFID 1 15 BY SIMILARITY.
FT DISULFID 3 11 BY SIMILARITY.
SQ SEQUENCE 21 AA; 2511 MW; 83A5DFB81D036AE2 CRC64;

Query Match 22.4%; Score 22; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQ 10
DB 15 CHQ 17

Search completed: April 23, 2003, 13:43:49
Job time : 5.75169 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:36:43 ; Search time 9.40449 Seconds
(without alignments)
183.999 Million cell updates/sec

Title: US-09-836-073-3
Perfect score: 98
Sequence: 1 AALQAKICHQIYFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 4984

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	32.7	22	2 PH1359	Ig heavy chain DJ
2	30	30.6	15	2 PH0789	T-cell receptor al
3	29	29.6	25	2 A60286	heat-stable serine
4	28	28.6	23	2 PH1725	Ig heavy chain V r
5	27	27.6	12	2 S57570	T cell receptor V-
6	27	27.6	14	2 S57569	T cell receptor V-
7	27	27.6	14	2 S57638	T cell receptor V-
8	27	27.6	22	2 PH1325	Ig heavy chain DJ
9	27	27.6	23	2 PH1691	Ig heavy chain V r
10	27	27.6	23	2 PH1724	Ig heavy chain V r
11	27	27.6	24	2 SA2780	relaxin - oranguta
12	27	27.6	24	2 PH1696	Ig heavy chain V r
13	27	27.6	25	2 C57001	endo-1,4-beta-xyla
14	26	26.5	12	2 PH1324	Ig heavy chain DJ
15	26	26.5	18	2 PH1368	Ig heavy chain DJ
16	26	26.5	22	2 I77373	gene N-ras protein
17	26	26.5	23	2 PH1682	Ig heavy chain V r
18	26	26.5	24	2 PT0258	Ig heavy chain CDR
19	25	25.5	12	2 G64003	hypothetical prote
20	25	25.5	21	2 S07968	T-cell receptor be
21	25	25.5	21	2 PH1730	Ig heavy chain V r
22	25	25.5	22	2 A40741	T-cell receptor be
23	25	25.5	22	2 PH1678	Ig heavy chain V r
24	25	25.5	22	2 PH1679	Ig heavy chain V r
25	25	25.5	23	2 PH1694	Ig heavy chain V r
26	25	25.5	23	2 PH1707	Ig heavy chain V r
27	25	25.5	23	2 PH1722	Ig heavy chain V r
28	25	25.5	23	2 PH1727	Ig heavy chain V r
29	25	25.5	23	2 PH1733	Ig heavy chain V r

ALIGNMENTS

RESULT 1

PH1359
Ig heavy chain DJ region (clone CI78-121) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PH1359
R:Wasserman, R.; Galilli, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor
A:Reference number: PH1302; MUID:93094761; PMID:1460419
A:Accession: PH1359
A:Molecule type: DNA
A:Residues: 1-22 <WAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 32.7%; Score 32; DB 2; Length 22;
Best Local Similarity 44.4%; Pred. No. 70;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 8 CHQIQYFYG 16
I::|::|
Db 6 CYENYYYG 14

RESULT 2

PH0789
T-cell receptor alpha chain (E22 V-alpha-4,delta-7R) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PH0789
R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A:Title: T cell receptor genes in a series of class I major histocompatibility comp
allelic exclusion and antigen-specific repertoire
A:Reference number: PH0746; MUID:92078846; PMID:1836010
A:Accession: PH0789
A:Molecule type: mRNA
A:Residues: 1-15 <CAS>
A:Cross-references: EMBL:X60894
A:Experimental source: T lymphocyte
C:Keywords: T-cell receptor

Query Match 30.6%; Score 30; DB 2; Length 15;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 9 HQIQYFYG 16
I::|::|
Db 8 HGLQYFYG 15

RESULT 3
A60286
Heat-stable serine proteinase (EC 3.4.21.-) - Thermomonospora fusca (strain YX) (fragment)
N:Alternate names: YX-proteinase
C:Species: Thermomonospora fusca
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Dec-1996
C:Accession: A60286
R:Kristjansson, M.M.; Kinsella, J.E.
Int. J. Pept. Protein Res. 36, 201-207, 1990
A:Title: Heat stable proteinase from Thermomonospora fusca. Characterization as a serine
A:Reference number: A60286; MUID:91107200; PMID:2132918
A:Accession: A60286
A:Molecule type: protein
A:Residues: 1-25 <NR>
C:Keywords: extracellular protein; hydrolase; serine proteinase

Query Match 29.6%; Score 29; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 YVFGQF 18
|||||
DB 10 YVFGNY 15

RESULT 4
PH1725
Ig heavy chain V region (clone GCC-5) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1725
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385
A:Accession: PH1725
A:Molecule type: mRNA
A:Residues: 1-23 <MCH>
A:Experimental source: B cell
A>Note: The authors translated the codon ACA for residue 13 as Ala
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 28.6%; Score 28; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 3.6e+02;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIQYFGQF 18
:|||||:
DB 12 ETRYIYGSY 20

RESULT 5
S57570
T cell receptor V-J junctional alpha chain region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57570
R:Burrows, S.R.; Sillins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaeet, V.P.
submitted to the EMBL Data Library, June 1995
A:Description: T cell receptor repertoire for a viral epitope in humans is diversified
A:Reference number: S57494
A:Accession: S57570
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-12 <BUR>
A:Cross-references: EMBL:249954; NID:9887488; PIDN:CAA90225.1; PID:9887489
C:Keywords: T-cell receptor

Query Match 27.6%; Score 27; DB 2; Length 12;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 QYFYG 16
|||||
DB 8 QYFYG 12

RESULT 6
S57569
T cell receptor V-J junctional alpha chain region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57569
R:Burrows, S.R.; Sillins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaeet, V.P.
submitted to the EMBL Data Library, June 1995
A:Description: T cell receptor repertoire for a viral epitope in humans is diversified
A:Reference number: S57494
A:Accession: S57569
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-14 <BUR>
A:Cross-references: EMBL:249955; NID:9887482; PIDN:CAA90226.1; PID:9887483
C:Keywords: T-cell receptor

Query Match 27.6%; Score 27; DB 2; Length 14;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 QYFYG 16
|||||
DB 10 QYFYG 14

RESULT 7
S57638
T cell receptor V-J junctional alpha chain region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57638
R:Burrows, S.R.; Sillins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaeet, V.P.
submitted to the EMBL Data Library, June 1995
A:Description: T cell receptor repertoire for a viral epitope in humans is diversified
A:Reference number: S57494
A:Accession: S57638
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-14 <BUR>
A:Cross-references: EMBL:249964; NID:9886676; PIDN:CAA90238.1; PID:9886677
C:Keywords: T-cell receptor

Query Match 27.6%; Score 27; DB 2; Length 14;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 QYFYG 16
|||||
DB 10 QYFYG 14

RESULT 8
PH1325
Ig heavy chain DJ region (clone C199-121) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PH1325
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor
A:Reference number: PH1302; MUID:93094761; PMID:1460419
A:Accession: PH1325
A:Molecule type: DNA
A:Residues: 1-22 <WAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 27.6%; Score 27; DB 2; Length 22;
Best Local Similarity 66.7%; Pred. No. 5.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 IOYIFG 16
| | | | |
Db 9 IHYYG 14

RESULT 9

PH1681
Ig heavy chain V region (clone NP-6-12) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1681
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385
A:Accession: PH1681
A:Molecule type: mRNA
A:Residues: 1-23 <MCH>
A:Experimental source: B cell
A:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 27.6%; Score 27; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 5.4e+02;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIOYFQGF 18
: | | | |
Db 12 EARYYGY 20

RESULT 10

PH1724
Ig heavy chain V region (clone GCC-4) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1724
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385
A:Accession: PH1724
A:Molecule type: mRNA
A:Residues: 1-23 <MCH>
A:Experimental source: B cell
A:Note: the authors translated the codon ACA for residue 13 as Ala
A:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 27.6%; Score 27; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 5.4e+02;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIOYFQGF 18
: | | | |
Db 12 ETRYGY 20

RESULT 11

S42780
relaxin - orangutan (fragment)
C:Species: Pongo pygmaeus (orangutan)
C:Date: 08-May-1995 #sequence_revision 26-Jul-1996 #text_change 16-Jul-1999
C:Accession: S42780
R:Evans, B.B.
submitted to the EMBL Data Library, November 1993
A:Reference number: S42776
A:Accession: S42780
A:Molecule type: DNA

A:Residues: 1-24 <EVA>
A:Cross-references: EMBL:227229; NID:9415990; PIDN:CAA81743.1; PID:9415991
C:Genetics:
A:Gene: rlxl
C:Superfamily: insulin
C:Keywords: disulfide bond; hormone
F:1-24/Domain: relaxin chain 1B (fragment) #status predicted <RXBI>

Query Match 27.6%; Score 27; DB 2; Length 24;
Best Local Similarity 55.6%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AALQAKICH 9
: | | | |
Db 4 SALANKCCH 12

RESULT 12

PH1696
Ig heavy chain V region (clone NP-7-9) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1696
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607; PMID:8315385
A:Accession: PH1696
A:Molecule type: mRNA
A:Residues: 1-24 <MCH>
A:Experimental source: B cell
A:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 27.6%; Score 27; DB 2; Length 24;
Best Local Similarity 42.9%; Pred. No. 5.6e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 10 QIOYFYG 16
: | | | |
Db 12 EVAYYIG 18

RESULT 13

C57001
endo-1,4-beta-xylanase (EC 3.2.1.8) 4 - Streptomyces roseiscleroticus (fragment)
C:Species: Streptomyces roseiscleroticus
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 23-May-1997
C:Accession: C57001
R:Grabski, A.C.; Forrester, I.T.; Patel, R.; Jeffries, T.W.
Protein Expr. Purif. 4, 120-129, 1993
A:Title: Characterization and N-terminal amino acid sequences of beta-(1-4)endoxylase
A:Reference number: A57001; MUID:93229899; PMID:8471845
A:Accession: C57001
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-25 <GRA>
A:Experimental source: strain NRRL B-11019
A:Note: sequence extracted from NCBI backbone (NCBIP:130009)
C:Function:
A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
A:Pathway: xylan degradation
C:Superfamily: Streptomyces endo-1,4-beta-xylanase A; Streptomyces endo-1,4-beta-xy
C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradati

Query Match 27.6%; Score 27; DB 2; Length 25;
Best Local Similarity 71.4%; Pred. No. 5.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIOYFYG 16
: | | | |
Db 11 QSGYFYG 17

RESULT 14

PHI324

Ig heavy chain DJ region (clone C510-100) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PHI324
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma
A:Reference number: PHI302; MUID:93094761; PMID:1460419
A:Accession: PHI324
A:Molecule type: DNA
A:Residues: 1-12 <WAS>
C:Keywords: heterotetramer; immunoglobulin

Query Match 26.5%; Score 26; DB 2; Length 12;
Best Local Similarity 80.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 13 YIFGQ 17

||:|

Db 6 YWQG 10

RESULT 15

PHI368

Ig heavy chain DJ region (clone C111-112) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PHI368
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma
A:Reference number: PHI302; MUID:93094761; PMID:1460419
A:Accession: PHI368
A:Molecule type: DNA
A:Residues: 1-18 <WAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 26.5%; Score 26; DB 2; Length 18;
Best Local Similarity 50.0%; Pred. No. 6.3e+02;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 11 IQYFG 16

:|:|

Db 5 MEYYG 10

Search completed: April 23, 2003, 13:48:53
Job time : 9.40449 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:43:09 ; Search time 9.91011 Seconds
(without alignments)
145.542 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98

Sequence: 1 AALQAKICHQIQYFGQF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 61792

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB_PEP.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB_PEP.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB_PEP.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB_PEP.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB_PEP.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB_PEP.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB_PEP.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB_PEP.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB_PEP.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB_PEP.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB_PEP.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB_PEP.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB_PEP.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB_PEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	100.0	18	9	US-09-836-073-3
2	87	88.8	18	9	US-09-836-073-1
3	87	88.8	18	9	US-09-836-073-14
4	83	84.7	17	9	US-09-836-073-13
5	81	82.7	18	9	US-09-836-073-9
6	79	80.6	18	9	US-09-836-073-11
7	79	80.6	18	9	US-09-836-073-12
8	78	79.6	18	9	US-09-836-073-10
9	76.5	78.1	19	9	US-09-836-073-16
10	75	76.5	18	9	US-09-836-073-2
11	75	76.5	18	9	US-09-836-073-4
12	72	73.5	18	9	US-09-836-073-7
13	71	72.4	18	9	US-09-836-073-8
14	65	66.3	18	9	US-09-836-073-15
15	63	64.3	18	9	US-09-836-073-5
16	56	57.1	18	9	US-09-836-073-6
17	45	45.9	16	9	US-09-836-073-19
18	39	39.8	18	9	US-09-836-073-17
19	37.5	38.3	23	9	US-10-011-585A-144

20 30.5 31.1 18 9 US-09-836-073-18
21 30 30.6 19 9 US-10-153-159-51
22 30 30.6 19 9 US-10-153-159-53
23 30 30.6 19 9 US-10-153-159-54
24 30 30.6 19 9 US-10-153-176-51
25 30 30.6 19 9 US-10-153-176-53
26 30 30.6 19 9 US-10-153-176-54
27 29 29.6 14 8 US-08-424-550B-455
28 29 29.6 19 9 US-10-153-176-6
29 29 29.6 19 9 US-10-153-176-6
30 29 29.6 20 9 US-09-986-480-269
31 28.5 29.1 23 9 US-09-795-515-9
32 28 28.6 19 9 US-10-153-159-49
33 28 28.6 19 9 US-10-153-159-59
34 28 28.6 19 9 US-10-153-176-49
35 28 28.6 19 9 US-10-153-176-59
36 28 28.6 19 9 US-09-880-748-2987
37 28 28.6 21 10 US-09-853-830-52
38 27 27.6 14 10 US-09-932-800-5
39 27 27.6 14 10 US-09-932-994-5
40 27 27.6 18 9 US-10-084-813-139
41 27 27.6 18 9 US-10-084-813-140
42 27 27.6 18 9 US-10-084-813-141
43 27 27.6 25 9 US-10-097-085-405
44 26 26.5 9 10 US-09-753-831-25
45 26 26.5 13 9 US-09-880-748-3091

ALIGNMENTS

RESULT 1
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; TYPE: PRT
; LENGTH: 18
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match 100.0%; Score 98; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18

DB 1 AALQAKICHQIQYFGQF 18

RESULT 2

US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-1

Query Match 88.8%; Score 87; DB 9; Length 18;
 Best Local Similarity 83.3%; Pred. No. 6.7e-08;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFQGF 18
 DB 1 AALEAKICHQIEYFGDF 18

RESULT 3
 US-09-836-073-14
 ; Sequence 14, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 220002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Bovine
 US-09-836-073-14

Query Match 88.8%; Score 87; DB 9; Length 18;
 Best Local Similarity 83.3%; Pred. No. 6.7e-08;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFQGF 18
 DB 1 AALEAKICHQIEYFGDF 18

RESULT 4
 US-09-836-073-13
 ; Sequence 13, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 220002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 17
 ; TYPE: PRT
 ; ORGANISM: Mouse
 US-09-836-073-13

Query Match 84.7%; Score 83; DB 9; Length 17;
 Best Local Similarity 82.4%; Pred. No. 2.9e-07;
 Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 AALQAKICHQIQYFQGF 18
 DB 1 AALEAKICHQIEYFGDF 17

RESULT 5
 US-09-836-073-9
 ; Sequence 9, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 220002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-9

Query Match 82.7%; Score 81; DB 9; Length 18;
 Best Local Similarity 87.5%; Pred. No. 6.4e-07;
 Matches 14; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFQGF 16
 DB 1 AALEAKICHQIEYFG 16

RESULT 6
 US-09-836-073-11
 ; Sequence 11, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 220002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-11

Query Match 80.6%; Score 79; DB 9; Length 18;
 Best Local Similarity 77.8%; Pred. No. 1.4e-06;
 Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFQGF 18
 DB 1 AALEAKICHQIEYFGDF 18

RESULT 7
 US-09-836-073-12

; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match 80.6%; Score 79; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 1.4e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
|||:|||||:||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 8

US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match 79.6%; Score 78; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 2e-06;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
|||:|||||:||||
Db 1 AALEAKICHQIEYQGF 18

RESULT 9

US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match 78.1%; Score 76.5; DB 9; Length 19;
Best Local Similarity 78.9%; Pred. No. 3.7e-06;
Matches 15; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALQAKICHQIQYFGQF 18
|||:|||||:||||
Db 1 AALEAKICHQIEYFGDF 19

RESULT 10

US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match 76.5%; Score 75; DB 9; Length 18;
Best Local Similarity 72.2%; Pred. No. 6.1e-06;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
|||:|||||:||||
Db 1 AALEAKICHQIEYFGDF 18

RESULT 11

US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match 76.5%; Score 75; DB 9; Length 18;
Best Local Similarity 80.0%; Pred. No. 6.1e-06;

Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 QAKICHQIQYFGQF 18
DB 4 EAKICHQIEYFGDF 18

RESULT 12

US-09-836-073-7
Sequence 7, Application US/09836073

Patent No. US20020173475A1

GENERAL INFORMATION:

APPLICANT: Dasgupta, Asim

APPLICANT: Das, S.

APPLICANT: Baidya, Narayan

TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

FILE REFERENCE: 22002054822

CURRENT APPLICATION NUMBER: US/09/836,073

CURRENT FILING DATE: 2002-10-24

PRIOR APPLICATION NUMBER: 09/316,630

PRIOR FILING DATE: 1999-05-21

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 18

TYPE: PRT

ORGANISM: Homo Sapiens

US-09-836-073-7

Query Match 73.5%; Score 72; DB 9; Length 18;

Best Local Similarity 81.2%; Pred. No. 1.9e-05;

Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFG 16

DB 1 AALEAKICHQIEYFG 16

RESULT 13

US-09-836-073-8

Sequence 8, Application US/09836073

Patent No. US20020173475A1

GENERAL INFORMATION:

APPLICANT: Dasgupta, Asim

APPLICANT: Das, S.

APPLICANT: Baidya, Narayan

TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

FILE REFERENCE: 22002054822

CURRENT APPLICATION NUMBER: US/09/836,073

CURRENT FILING DATE: 2002-10-24

PRIOR APPLICATION NUMBER: 09/316,630

PRIOR FILING DATE: 1999-05-21

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 8

LENGTH: 18

TYPE: PRT

ORGANISM: Homo Sapiens

US-09-836-073-8

Query Match 72.4%; Score 71; DB 9; Length 18;

Best Local Similarity 72.2%; Pred. No. 2.7e-05;

Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFG 18

DB 1 AALEAKICHQIEYFG 18

RESULT 14

US-09-836-073-15

Sequence 15, Application US/09836073

Patent No. US20020173475A1

GENERAL INFORMATION:

APPLICANT: Dasgupta, Asim

APPLICANT: Das, S.

APPLICANT: Baidya, Narayan

TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

FILE REFERENCE: 22002054822

CURRENT APPLICATION NUMBER: US/09/836,073

CURRENT FILING DATE: 2002-10-24

PRIOR APPLICATION NUMBER: 09/316,630

PRIOR FILING DATE: 1999-05-21

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 15

LENGTH: 18

TYPE: PRT

ORGANISM: Xenopus

US-09-836-073-15

Query Match 66.3%; Score 65; DB 9; Length 18;

Best Local Similarity 68.8%; Pred. No. 0.00026;

Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 LQAKICHQIQYFGQF 18

DB 3 LDTKICEIQIEYFGDF 18

RESULT 15

US-09-836-073-5

Sequence 5, Application US/09836073

Patent No. US20020173475A1

GENERAL INFORMATION:

APPLICANT: Dasgupta, Asim

APPLICANT: Das, S.

APPLICANT: Baidya, Narayan

TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

FILE REFERENCE: 22002054822

CURRENT APPLICATION NUMBER: US/09/836,073

CURRENT FILING DATE: 2002-10-24

PRIOR APPLICATION NUMBER: 09/316,630

PRIOR FILING DATE: 1999-05-21

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5

LENGTH: 18

TYPE: PRT

ORGANISM: Homo Sapiens

US-09-836-073-5

Query Match 64.3%; Score 63; DB 9; Length 18;

Best Local Similarity 66.7%; Pred. No. 0.00056;

Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 QAKICHQIQYFGQF 18

DB 4 EKOCHQIEYFGDF 18

Search completed: April 23, 2003, 13:52:08

Job time : 9.91011 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:38:28 ; Search time 8.49438 seconds
(without alignments)
62.349 Million cell updates/sec

Title: US-09-836-073-3

Perfect score: 98
Sequence: 1 AALQAKICHQIQYFGQF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 147762

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCRUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	87	88.8	18	4	US-09-316-630-3
2	87	88.8	18	4	US-09-316-630-4
3	33	33.7	14	2	US-08-637-759B-113
4	33	33.7	14	3	US-08-871-355A-113
5	33	33.7	14	4	US-09-201-945-113
6	31	31.6	20	4	US-08-505-250-8
7	31	31.6	20	4	US-08-505-250-8
8	30	30.6	11	2	US-08-211-312-9
9	30	30.6	11	3	US-08-472-285-9
10	30	30.6	11	4	US-08-472-285-9
11	29	29.6	14	4	US-08-469-260A-455
12	28.5	29.1	21	1	US-08-447-411-38
13	28.5	29.1	23	2	US-08-303-569B-9
14	28.5	29.1	23	4	US-08-737-629-8
15	28	28.6	15	1	US-08-221-581-1
16	28	28.6	15	5	PCT-US95-04018-72
17	28	28.6	21	2	US-08-825-349-1
18	28	28.6	22	6	5281520-43
19	27	27.6	7	1	US-08-166-930-15
20	27	27.6	7	2	US-08-727-045A-15
21	27	27.6	7	4	US-09-408-172-15
22	27	27.6	9	4	US-08-747-599A-12
23	27	27.6	12	4	US-08-737-841-16
24	27	27.6	14	2	US-08-726-464B-47
25	27	27.6	17	1	US-08-554-612C-44
26	27	27.6	19	6	5464756-40
27	27	27.6	24	1	US-08-443-568B-4

Sequence 2, Appl1
Sequence 89, Appl1
Sequence 2, Appl1
Sequence 2, Appl1
Sequence 4, Appl1
Patent No. 5464756
Sequence 4, Appl1
Sequence 72, Appl1
Sequence 73, Appl1
Sequence 4, Appl1
Sequence 72, Appl1
Sequence 73, Appl1
Patent No. 5464218
Sequence 94, Appl1
Sequence 472, Appl1
Sequence 29, Appl1
Sequence 254, Appl1
Sequence 29, Appl1

28 27 27.6 24 2 US-08-483-476-2
29 27 27.6 24 2 US-08-353-476-89
30 27 27.6 24 2 US-08-484-219-2
31 27 27.6 24 4 US-09-158-706-2
32 27 27.6 24 5 PCT-US94-06997-4
33 27 27.6 24 6 5464756-2
34 27 27.6 25 1 US-08-453-289-4
35 27 27.6 25 1 US-08-614-935-72
36 27 27.6 25 1 US-08-614-935-73
37 27 27.6 25 2 US-08-574-086-4
38 27 27.6 25 3 US-09-130-287-72
39 27 27.6 25 3 US-09-130-287-73
40 27 27.6 25 6 5424218-1
41 26 26.5 14 2 US-08-433-133-94
42 26 26.5 14 4 US-08-469-260A-472
43 26 26.5 15 4 US-08-743-168B-29
44 26 26.5 15 4 US-09-009-953-254
45 26 26.5 15 5 PCT-US96-10435-29

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Samittra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match 88.8%; Score 87; DB 4; Length 18;
Best Local Similarity 83.3%; Pred. No. 4.9e-08;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18

Db 1 AALQAKICHQIEYFGDF 18

RESULT 2

US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Samittra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

; PRIOR FILING DATE: 1997-10-06
 ; PRIOR APPLICATION NUMBER: 08/321,427
 ; PRIOR FILING DATE: 1994-10-11
 ; PRIOR APPLICATION NUMBER: 60/086,527
 ; PRIOR FILING DATE: 1998-05-22
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.1.
 ; SEQ ID NO 4
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
 ; OTHER INFORMATION: this peptide is biotinylated
 US-09-316-630-4

Query Match 88.8%; Score 87; DB 4; Length 18;
 Best Local Similarity 83.3%; Pred. No. 4.9e-08;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
 DB 1 AALEAKICHQIYFGDF 18

RESULT 3
 US-08-637-759B-113
 ; Sequence 113, Application US/08637759B
 ; Patent No. 5876931
 ; GENERAL INFORMATION:
 ; APPLICANT: David William Holden
 ; TITLE OF INVENTION: Identification of Genes
 ; NUMBER OF SEQUENCES: 501
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Patrea L. Pabst
 ; STREET: 2800 One Atlantic Center
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: USA
 ; ZIP: 30309-3450
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/637,759B
 ; FILING DATE: 03-MAY-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/GB95/02875
 ; FILING DATE: 11-DEC-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pabst, Patrea L.
 ; REGISTRATION NUMBER: 31,284
 ; REFERENCE/DOCKET NUMBER: RPMS 101
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (404) 873-8794
 ; TELEFAX: (404) 873-8795
 ; INFORMATION FOR SEQ ID NO: 113:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 14 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO

US-08-637-759B-113
 Query Match 33.7%; Score 33; DB 2; Length 14;
 Best Local Similarity 41.7%; Pred. No. 24;
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 ICHQIQYFGQF 18
 DB 1 VCNNIQYSGHY 12

Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 7 ICHQIQYFGQF 18
 DB 1 VCNNIQYSGHY 12

RESULT 4
 US-08-871-355A-113
 ; Sequence 113, Application US/08871355A
 ; Patent No. 6015669
 ; GENERAL INFORMATION:
 ; APPLICANT: David William Holden
 ; TITLE OF INVENTION: Identification of Genes
 ; NUMBER OF SEQUENCES: 501
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Patrea L. Pabst
 ; STREET: 2800 One Atlantic Center
 ; CITY: Atlanta
 ; STATE: Georgia
 ; COUNTRY: USA
 ; ZIP: 30309-3450
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/871,355A
 ; FILING DATE: 09-JUN-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/GB95/02875
 ; FILING DATE: 11-DEC-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pabst, Patrea L.
 ; REGISTRATION NUMBER: 31,284
 ; REFERENCE/DOCKET NUMBER: RPMS 101
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (404) 873-8794
 ; TELEFAX: (404) 873-8795
 ; INFORMATION FOR SEQ ID NO: 113:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 14 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO

US-08-871-355A-113
 Query Match 33.7%; Score 33; DB 3; Length 14;
 Best Local Similarity 41.7%; Pred. No. 24;
 Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 ICHQIQYFGQF 18
 DB 1 VCNNIQYSGHY 12

RESULT 5
 US-09-201-945-113
 ; Sequence 113, Application US/09201945
 ; Patent No. 6342215
 ; GENERAL INFORMATION:
 ; APPLICANT: David William Holden
 ; TITLE OF INVENTION: Identification of Genes
 ; NUMBER OF SEQUENCES: 501
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Patrea L. Pabst
 ; STREET: 2800 One Atlantic Center

us-09-836-073-3.lim25.ra

Thu Apr 24 08:54:08 2003

STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/09/201,945
APPLICATION NUMBER: US/09/201,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: 08/637,759
APPLICATION NUMBER: 08/637,759
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: RPS 101
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-201-945-113

Query Match 33.7%; Score 33; DB 4; Length 14;
Best Local Similarity 41.7%; Pred. No. 24;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 7 ICHQIOYFGQF 18
|:|:|:|:|:
Db 1 VCNNIQYKRGY 12

RESULT 6
US-08-505-250-8
; Sequence 8, Application US/08505250
; Patent No. 6183983
; GENERAL INFORMATION:
; APPLICANT: Sato, Haruya
; APPLICANT: Yamamoto, Keiji
; APPLICANT: Suzuki, Kokichi
; APPLICANT: Ikeda, Masahiro
; APPLICANT: Sakagami, Masahiro
; APPLICANT: Taniguchi, Makoto
; TITLE OF INVENTION: PROTEIN MODIFICATION METHOD
; FILE REFERENCE: 110-511
; CURRENT APPLICATION NUMBER: US/08/505,250
; CURRENT FILING DATE: 1995-11-29
; EARLIER APPLICATION NUMBER: PCT/JP95/00298
; EARLIER FILING DATE: 1995-02-27
; EARLIER APPLICATION NUMBER: JP 198187/94
; EARLIER FILING DATE: 1994-08-23
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-08-505-250-8

Query Match 31.6%; Score 31; DB 4; Length 20;
Best Local Similarity 55.6%; Pred. No. 73;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIQY 14
|:|:|:|:|:
Db 3 KMCPLQY 11

RESULT 7
US-08-505-250-8
; Sequence 8, Application US/08505250
; Patent No. 6322996
; GENERAL INFORMATION:
; APPLICANT: Sato, Haruya
; APPLICANT: Yamamoto, Keiji
; APPLICANT: Suzuki, Kokichi
; APPLICANT: Ikeda, Masahiro
; APPLICANT: Sakagami, Masahiro
; APPLICANT: Taniguchi, Makoto
; TITLE OF INVENTION: PROTEIN MODIFICATION METHOD
; FILE REFERENCE: 110-511
; CURRENT APPLICATION NUMBER: US/08/505,250
; CURRENT FILING DATE: 1995-11-29
; PRIOR APPLICATION NUMBER: PCT/JP95/00298
; PRIOR FILING DATE: 1995-02-27
; PRIOR APPLICATION NUMBER: JP 198187/94
; PRIOR FILING DATE: 1994-08-23
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-08-505-250-8

Query Match 31.6%; Score 31; DB 4; Length 20;
Best Local Similarity 55.6%; Pred. No. 73;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 6 KICHOIQY 14
|:|:|:|:|:
Db 3 KMCPLQY 11

RESULT 8
US-08-211-312-9
; Sequence 9, Application US/08211312
; Patent No. 5986051
; GENERAL INFORMATION:
; APPLICANT: LABIGNE, AGNES
; APPLICANT: CUSSAC, VALERIE
; APPLICANT: FERRERO, RICHARD
; TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
; TITLE OF INVENTION: FOR THE REGULATION AND MATURATION OF UREASE AND THEIR USE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
US-08-211-312-9

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/211,312
;; FILING DATE: 01-JUL-1994
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: FR 91 12198
;; FILING DATE: 03-OCT-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/FR92/00921
;; FILING DATE: 02-OCT-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Oblon, No. 5986051man F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 660-075-0XPCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 413-3000
;; TELEFAX: (703) 413-2220
;; TELEX: 248855 OPAT UR
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 11 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
US-08-211-312-9

Query Match 30.6%; Score 30; DB 2; Length 11;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 AKICHOI 11
||||:|
Db 1 AKICYEI 7

RESULT 9
US-08-472-285-9
;; Sequence 9, Application US/08472285
;; Patent No. 6027878
;; GENERAL INFORMATION:
;; APPLICANT: LABIGNE, AGNES
;; APPLICANT: CUSSAC, VALERIE
;; APPLICANT: FERRERO, RICHARD
;; TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
;; FOR THE REGULATION AND MATURATION OF UREASE AND THEIR USE
;; NUMBER OF SEQUENCES: 12
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
;; STREET: 1755 S. Jefferson Davis Highway, Suite 400
;; CITY: Arlington
;; STATE: Virginia
;; COUNTRY: U.S.A.
;; ZIP: 22202
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/211,312
;; FILING DATE: 01-JUL-1994
;; APPLICATION NUMBER: FR 91 12198
;; FILING DATE: 03-OCT-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/FR92/00921
;; FILING DATE: 02-OCT-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Oblon, No. 6027878man F.

;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 660-075-0XPCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 413-3000
;; TELEFAX: (703) 413-2220
;; TELEX: 248855 OPAT UR
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 11 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
US-08-472-285-9

Query Match 30.6%; Score 30; DB 3; Length 11;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 AKICHOI 11
||||:|
Db 1 AKICYEI 7

RESULT 10
US-08-472-929-9
;; Sequence 9, Application US/08472929
;; Patent No. 6271017
;; GENERAL INFORMATION:
;; APPLICANT: LABIGNE, AGNES
;; APPLICANT: CUSSAC, VALERIE
;; APPLICANT: FERRERO, RICHARD
;; TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
;; FOR THE REGULATION AND MATURATION OF UREASE AND THEIR USE
;; NUMBER OF SEQUENCES: 12
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
;; STREET: 1755 S. Jefferson Davis Highway, Suite 400
;; CITY: Arlington
;; STATE: Virginia
;; COUNTRY: U.S.A.
;; ZIP: 22202
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; FILING DATE: US/08/472,929
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/211,312
;; FILING DATE:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/FR92/00921
;; FILING DATE: 02-OCT-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Oblon, No. 6271017man F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 660-075-0XPCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 413-3000
;; TELEFAX: (703) 413-2220
;; TELEX: 248855 OPAT UR
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 11 amino acids
;; TYPE: amino acid
;; TOPOLOGY: unknown
;; MOLECULE TYPE: peptide
US-08-472-929-9

Thu Apr 24 08:54:08 2003

Query Match 30.6%; Score 30; DB 4; Length 11;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 AKICHOI 11
| | | | |
DB 1 AKICYEI 7

RESULT 11
US-08-469-260A-455
; Sequence 455, Application US/08469260A
; Patent No. 6451578
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAMSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSAHAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,260A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: POREBSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 455:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-469-260A-455

Query Match 29.6%; Score 29; DB 4; Length 14;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 AKICHO 10
| | | | |
DB 3 AMICHO 8

RESULT 12
US-08-447-411-38
; Sequence 38, Application US/08447411

; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Naja naja
; US-08-447-411-38
; Query Match 29.1%; Score 28.5; DB 1; Length 21;
; Best Local Similarity 43.8%; Pred. No. 2e+02; 3; Indels 3; Gaps 1;
; Matches 7; Conservative 3; Mismatches 3;
QY 2 ALQAKICHOIQYVFGQ 17
| | | | |
DB 1 ALRLKICTR---YLGE 13
; RESULT 13
US-08-303-569B-9
; Sequence 9, Application US/08303569B
; Patent No. 5859205
; GENERAL INFORMATION:
; APPLICANT: Adair, John R.
; APPLICANT: Athwal, Diljeet S.
; APPLICANT: Entage, John S.
; TITLE OF INVENTION: Humanised Antibodies
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5859205r1s
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

TYPE: amino acid

PT New yeast inhibitory peptide useful for inhibiting viral protein
 translation and replication -
 XX
 PS Claim 5; Page 57; 81pp; English.
 XX
 CC This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (IRNA see AAZ45200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rhadovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.

XX Sequence 18 AA;

Query Match 88.8%; Score 87; DB 21; Length 18;
 Best Local Similarity 83.3%; Pred. No. 3.1e-07;
 Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALQAKICHQIQYFGQF 18
 |||:|||||:|||||
 Db 1 AALEAKICHQIEYFGDF 18

RESULT 2

AA43394
 ID AAR43394 standard; peptide; 21 AA.

XX AAR43394;

DT 12-MAY-1994 (first entry)

DE La/SSB epitope 17.

KW Linear: epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

XX Homo sapiens.

PN W09321223-A.

XX 28-OCT-1993.

PF 13-APR-1993; 93WO-US03484.

PR 13-APR-1992; 92US-0867819.

XX (OKLA) UNIV OKLAHOMA STATE.

PI Harley JB;

DR WPI; 1993-351658/44.

XX New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus

XX Claim 1; Page 30; 43pp; English.

XX The sequences given in AAR43391-562 are linear epitopes which are

CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polypyridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.

XX Sequence 21 AA;

Query Match 65.3%; Score 64; DB 14; Length 21;
 Best Local Similarity 83.3%; Pred. No. 0.0018;
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIQYFGQF 18
 |||||:|||||
 Db 1 ICHQIEYFGDF 12

RESULT 3

ABB79195

ID ABB79195 standard; Protein; 23 AA.

XX ABB79195;

DT 08-AUG-2002 (first entry)

DE Human prostate specific protein sequence SEQ ID NO:144.

XX Human; prostate specific gene; prostate specific protein; PSP;
 KW prostate cancer.

XX Homo sapiens.

PN W0200236808-A2.

XX 10-MAY-2002.

XX 05-NOV-2001; 2001WO-US47283.

XX 03-NOV-2000; 2000US-245740P.

XX (DIAD-) DIADEXUS INC.

XX Sun Y, Recipon H, Chen S, Liu C;

XX WPI; 2002-471506/50.

XX New prostate-specific nucleic acids and polypeptides, useful for
 PT identifying, diagnosing, monitoring, staging, imaging, and treating
 PT prostate cancer and non-cancerous disease states in prostate tissue

XX Claim 11; Page 218; 254pp; English.

XX ABN87650 to ABN87789 represent human prostate-specific nucleic acids (I),
 CC and ABN879192 to ABN879295 represent human prostate-specific proteins (II)
 CC from the present invention. (I) and (II) have cytostatic activity. (I)
 CC can be used in gene therapy. The prostate-specific nucleic acids,
 CC polypeptides and compositions from the present invention can be used for
 CC identifying, diagnosing, monitoring, staging, imaging, and treating
 CC prostate cancer and non-cancerous disease states in prostate tissue; for

CC Identifying prostate tissue; for monitoring, identifying and/or designing
 CC agonists and antagonists of the polypeptides; in gene therapy; in
 CC producing transgenic animals and cells; for producing engineered prostate
 CC tissue for treatment and research; and as elements in an array or
 CC computer program for pattern recognition of prostate disorders. The
 CC nucleic acids may be used as hybridisation probes to detect, characterise
 CC and quantify hybridising nucleic acids in, and isolate hybridising
 CC nucleic acids from, both genomic and transcript-derived nucleic acid
 CC samples.

XX
 SQ Sequence 23 AA;
 Query Match 38.3%; Score 37.5; DB 23; Length 23;
 Best Local Similarity 53.3%; Pred. No. 35;
 Matches 8; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 ALQAKICHQI-QYFF 15
 III III :|||
 Db 6 ALQTPICHTTGKPYF 20

RESULT 4
 ABB84046
 ID ABB84046 standard; peptide; 10 AA.

XX AC ABB84046;

XX DT 21-AUG-2002 (first entry)

XX DE Transglutaminase inhibitory peptide cr type #16.

XX KW Transglutaminase inhibitor; Factor XIII inhibitor; XIIIa inhibitor;
 KW ophthalmological; antiinflammatory; antirheumatic; antiarthritic;
 KW thrombolytic; neuroprotective; nootropic; antiseborrheic; dermatological;
 KW cyostatic; anti-HIV; antipsoriatic; cataract; inflammatory disease;
 KW arthritis; thrombosis; Alzheimer's disease; Huntington's chorea;
 KW acne; cancer; HIV infection; psoriasis.

XX OS Unidentified.

XX PN WO200236798-A2.

XX PD 10-MAY-2002.

XX PF 02-NOV-2001; 2001WO-EP12727.

XX PR 03-NOV-2000; 2000DE-1054687.

XX PA (NZYM-) N ZYME BIOPEC GMBH.

XX PI Fuchsbaauer H, Pasternack R, Zotzel J;

XX DR WPI; 2002-444364/47.

XX PT New amino acid or peptide derivatives or analogs, are selective
 PT transglutaminase inhibitors useful e.g. for treating cataract,
 PT inflammatory diseases, rheumatoid arthritis, thrombosis, Alzheimer's
 PT disease and cancer

XX PS Disclosure; Page 13; 44pp; German.

XX CC This invention describes novel amino acid or peptide derivatives or
 CC analogues (I), containing a modified side-chain (e.g. containing a formyl
 CC group) which are transglutaminase inhibitors and Factor XIII/XIIIa
 CC inhibitors. The products of the invention have ophthalmological,
 CC antiinflammatory, antirheumatic, antiarthritic, thrombolytic,
 CC neuroprotective, nootropic, antiseborrheic, dermatological, cyostatic,
 CC anti-HIV and antipsoriatic activity. (I) are transglutaminase inhibitors,
 CC especially inhibitors of crosslinking of proteins or peptides
 CC (specifically fibrin and/or alpha₂-plasmin inhibitor), incorporation of
 CC primary amines in proteins and peptides, hydrolysis of the
 CC gamma-carboxamido group of glutamine residues bound in proteins or
 CC peptides, blood factor XIII/XIIIa and mammalian, human, tissue, liver,

CC brain, eye lens, keratinocyte, epidermal, prostate, plant, parasitic
 CC and/or bacterial transglutaminases. The products of the invention can be
 CC used for treating cataract, inflammatory diseases, rheumatoid arthritis,
 CC chronic arthritis, thrombosis, Alzheimer's disease, Huntington's chorea,
 CC acne, cancer (by induction of apoptosis), HIV infections and psoriasis.
 CC (I) are targeted and specific transglutaminase inhibitors, which can
 CC inhibit a specific type of transglutaminase in the human or animal body
 CC without affecting other transglutaminases. ABB84001-ABB84049 represent
 CC transglutaminase inhibitors described in the method of the invention.

XX
 SQ Sequence 10 AA;

Query Match 34.7%; Score 34; DB 23; Length 10;
 Best Local Similarity 85.7%; Pred. No. 54;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 QAKICHQ 10
 II IIII
 Db 3 QAPICHQ 9

RESULT 5
 ABG66356
 ID ABG66356 standard; Peptide; 21 AA.

XX AC ABG66356;

XX DT 30-AUG-2002 (first entry)

XX DE IgE Fcepsilon RI binding peptide IGE120 #20.

XX KW IgE receptor; Immunoglobulin; FcepsilonRI; antagonist; phage display;
 KW protein co-ordinate data; IgE-mediated disease; allergic rhinitis;
 KW asthma; allergic asthma; atopic dermatitis; urticaria-angioedema;
 KW parasitic infection; IgE myeloma; immune-related disorder;
 KW inflammatory disorder; diabetes mellitus; reperfusion injury; stroke;
 KW IgE-mediated gastrointestinal inflammatory disease; burn;
 KW immune rejection of graft; myocardial infarction; atherosclerosis;
 KW acute lung injury; haemorrhagic shock; septic shock;
 KW acute tubular necrosis; endometriosis; degenerative joint disease;
 KW pancreatitis.

XX OS Synthetic.

XX PN WO200226781-A2.

XX PD 04-APR-2002.

XX PF 26-SEP-2001; 2001WO-US30289.

XX PR 26-SEP-2000; 2000US-235353p.

XX PR 23-MAR-2001; 2001US-278540p.

XX PA (GETH) GENENTECH INC.

XX PI Lowman HB, Reynolds ME, Nakamura GR, Starovasnik MA;

XX DR WPI; 2002-444016/47.

XX PT A peptide useful for treating a IgE-mediated disease or disorder in a
 PT host e.g. allergic rhinitis, asthma, which competes with immunoglobulin
 PT E for binding to high affinity IgE receptor in an in vitro assay

XX PS Example 8; Fig 2; 328pp; English.

XX CC The invention relates to a peptide which competes with immunoglobulin
 CC (Ig) E 134 comprising a sequence (SI), for binding the high affinity IgE
 CC receptor (FcepsilonRI) in an in vitro assay and having a formula given in
 CC the specification. Also included are a fusion protein comprising the
 CC peptide, a pharmaceutical composition (C) comprising the peptide,
 CC designing a compound that mimics the three-dimensional surface
 CC structure of the peptide, a compound with a solvent accessible surface
 CC that mimics the solvent accessible surface defined by the side chains of

CC residues (R) Pro4, Phe6, Pro16, Cys3, Cys7, Cys15 and Cys19 of IgE134,
 CC a peptide with structural coordinates as given in the specification,
 CC selecting a peptide mimetic which binds to FcεpsilonRI and blocks
 CC binding of IgE and a peptide mimetic which mimics the coordinates of
 CC IgE134 residues (R). (C) is useful for inhibiting the binding of IgE to
 CC high affinity IgE receptor (FcεpsilonRI). Peptides of the formula given
 CC in the specification are useful for inhibiting the binding of an IgE to
 CC high affinity IgE receptor. The peptide is useful for selecting a
 CC molecule which blocks the interaction of IgE with high affinity IgE
 CC receptor. The peptide is also useful for inhibiting the activation of
 CC high affinity IgE receptor. The peptide is useful for treating an IgE-
 CC mediated disease or disorder in a host. (C) is useful in research,
 CC diagnostic, therapeutic and prophylactic methods. The peptide is also
 CC useful for inhibiting IgE-mediated or associated processes such as IgE-
 CC dependent activation and degranulation of mast cells and basophils, as
 CC well as consequent release of inflammatory mediators such as histamine.
 CC (C) is useful for treating allergic rhinitis, asthma (e.g. allergic
 CC asthma), atopic dermatitis, urticaria-angioedema, parasitic infection,
 CC IgE myeloma, immune-related disorders, inflammatory disorders, diabetes
 CC mellitus, IgE-mediated gastrointestinal inflammatory disease, immune
 CC rejection of grafts, reperfusion injury, stroke, myocardial infarction,
 CC atherosclerosis, acute lung injury, haemorrhagic shock, burn, septic
 CC shock, acute tubular necrosis, endometriosis, degenerative joint disease
 CC and pancreatitis. The present sequence is a peptide of the invention
 CC expressed from a phage display library.

XX Sequence 21 AA;

Query Match 31.6%; Score 31; DB 23; Length 21;

Best Local Similarity 40.0%; Pred. No. 3.5e+02;

Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIOYFG 16

Db :|:|:|:|:|
 2 VCPRLCYWFG 11

RESULT 6

AAU92267
 ID AAU92267 standard; peptide; 9 AA.

XX AC AAU92267;

XX DT 02-JUL-2002 (first entry)

XX DE PHORI-F5D6 peptide #5 tested for HLA binding.

XX KW Human; PHORI-All; PHORI-F5D6; prostate cancer; cytostatic.

XX OS Homo sapiens.

XX PN WO200214501-A2.

XX PD 21-FEB-2002.

XX PF 17-AUG-2001; 2001WO-US25862.

XX PR 17-AUG-2000; 2000US-226241P.

XX PA (AGEN-) AGENSYS INC.

XX PI Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;

XX PI Jakobovits A;

XX XX WPI; 2002-269193/31.

XX Monitoring PHORI-All/PHORI-F5D6 gene products for monitoring presence
 PT of cancer in subject, by determining status of PHORI-All/PHORI-F5D6
 PT gene products in tissue sample from subject and comparing it to normal
 PT sample -

XX Claim 49; Page 175; 250pp; English.

XX

CC The present invention relates to the isolation of novel human genes
 CC designated PHORI-All and PHORI-F5D6 and their encoded proteins. The
 CC gene encoding PHORI-All maps to chromosome 1q23, and the gene encoding
 CC PHORI-F5D6 maps to chromosome 7q33-q35. The PHORI-All and PHORI-F5D6
 CC polynucleotide and polypeptide sequences are useful in diagnostic and
 CC therapeutic methods, and compositions for various cancers such as
 CC prostate cancer. The sequences are useful for inhibiting the growth of
 CC cancer cells that express PHORI-All or PHORI-F5D6 and for treating
 CC cancer. The PHORI-All or PHORI-F5D6 polypeptide or a fragment thereof
 CC can be used to elicit an immune response. AAU91563-AAU92962 represent
 CC PHORI-All or PHORI-F5D6 peptides tested for HLA binding.

XX Sequence 9 AA;

Query Match 30.6%; Score 30; DB 23; Length 9;

Best Local Similarity 57.1%; Pred. No. 7.8e+05;

Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIOY 13

Db :|:|:|:|:|
 3 ICHPLRY 9

RESULT 7

AAU92288
 ID AAU92288 standard; peptide; 9 AA.

XX AC AAU92288;

XX DT 02-JUL-2002 (first entry)

XX DE PHORI-F5D6 peptide #26 tested for HLA binding.

XX KW Human; PHORI-All; PHORI-F5D6; prostate cancer; cytostatic.

XX OS Homo sapiens.

XX PN WO200214501-A2.

XX PD 21-FEB-2002.

XX PF 17-AUG-2001; 2001WO-US25862.

XX PR 17-AUG-2000; 2000US-226241P.

XX PA (AGEN-) AGENSYS INC.

XX PI Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
 PI Jakobovits A;

XX XX WPI; 2002-269193/31.

XX Monitoring PHORI-All/PHORI-F5D6 gene products for monitoring presence
 PT of cancer in subject, by determining status of PHORI-All/PHORI-F5D6
 PT gene products in tissue sample from subject and comparing it to normal
 PT sample -

XX Claim 49; Page 175; 250pp; English.

XX The present invention relates to the isolation of novel human genes
 CC designated PHORI-All and PHORI-F5D6 and their encoded proteins. The
 CC gene encoding PHORI-All maps to chromosome 1q23, and the gene encoding
 CC PHORI-F5D6 maps to chromosome 7q33-q35. The PHORI-All and PHORI-F5D6
 CC polynucleotide and polypeptide sequences are useful in diagnostic and
 CC therapeutic methods, and compositions for various cancers such as
 CC prostate cancer. The sequences are useful for inhibiting the growth of
 CC cancer cells that express PHORI-All or PHORI-F5D6 and for treating
 CC cancer. The PHORI-All or PHORI-F5D6 polypeptide or a fragment thereof
 CC can be used to elicit an immune response. AAU91563-AAU92962 represent
 CC PHORI-All or PHORI-F5D6 peptides tested for HLA binding.

XX Sequence 9 AA;

XX

Query Match 30.6%; Score 30; DB 23; Length 9;
Best Local Similarity 57.1%; Pred. No. 7.8e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIOY 13
||| :||
Db 1 ICHPLRY 7

RESULT 8
AAU92398
ID AAU92398 standard; peptide; 9 AA.

XX AC AAU92398;
XX DT 02-JUL-2002 (first entry)
XX DE PHORI-F5D6 peptide #136 tested for HLA binding.
XX KW Human; PHORI-All; PHORI-F5D6; prostate cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO200214501-A2.
XX PD 21-FEB-2002.
XX PF 17-AUG-2001; 2001WO-US25862.
XX PR 17-AUG-2000; 2000US-226241P.
XX PA (AGEN-) AGENSYS INC.
XX PI Hubert RS, Raitano AB, Faris M, Challita-Eld PM, Ge W;
XX PJ Jakobovits A;
XX DR WPI; 2002-269193/31.

XX PN WO200214501-A2.
XX PD 21-FEB-2002.
XX PF 17-AUG-2001; 2001WO-US25862.
XX PR 17-AUG-2000; 2000US-226241P.
XX PA (AGEN-) AGENSYS INC.
XX PI Hubert RS, Raitano AB, Faris M, Challita-Eld PM, Ge W;
XX PJ Jakobovits A;
XX DR WPI; 2002-269193/31.

XX PT Monitoring PHORI-All/PHORI-F5D6 gene products for monitoring presence
XX of cancer in subject, by determining status of PHORI-All/PHORI-F5D6
XX gene products in tissue sample from subject and comparing it to normal
XX sample -
XX Claim 49; Page 180; 250pp; English.
XX CC The present invention relates to the isolation of novel human genes
XX designated PHORI-All and PHORI-F5D6 and their encoded proteins. The
XX gene encoding PHORI-All maps to chromosome 1q23, and the gene encoding
XX PHORI-F5D6 maps to chromosome 7q33-q35. The PHORI-All and PHORI-F5D6
XX polynucleotide and polypeptide sequences are useful in diagnostic and
XX therapeutic methods, and compositions for various cancers such as
XX prostate cancer. The sequences are useful for inhibiting the growth of
XX cancer cells that express PHORI-All or PHORI-F5D6 and for treating
XX cancer. The PHORI-All or PHORI-F5D6 polypeptide or a fragment thereof
XX can be used to elicit an immune response. AAU91563-AAU92962 represent
XX PHORI-All or PHORI-F5D6 peptides tested for HLA binding.

XX Sequence 9 AA;

Query Match 30.6%; Score 30; DB 23; Length 9;
Best Local Similarity 57.1%; Pred. No. 7.8e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIOY 13
||| :||
Db 2 ICHPLRY 8

RESULT 9
AAU92773
ID AAU92773 standard; peptide; 9 AA.

XX AC AAU92773;
XX DT 02-JUL-2002 (first entry)
XX DE PHORI-F5D6 peptide #613 tested for HLA binding.
XX KW Human; PHORI-All; PHORI-F5D6; prostate cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO200214501-A2.
XX PD 21-FEB-2002.
XX PF 17-AUG-2001; 2001WO-US25862.

XX DT 02-JUL-2002 (first entry)
XX DE PHORI-F5D6 peptide #511 tested for HLA binding.
XX KW Human; PHORI-All; PHORI-F5D6; prostate cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO200214501-A2.
XX PD 21-FEB-2002.
XX PF 17-AUG-2001; 2001WO-US25862.
XX PR 17-AUG-2000; 2000US-226241P.
XX PA (AGEN-) AGENSYS INC.
XX PI Hubert RS, Raitano AB, Faris M, Challita-Eld PM, Ge W;
XX PJ Jakobovits A;
XX DR WPI; 2002-269193/31.

XX PT Monitoring PHORI-All/PHORI-F5D6 gene products for monitoring presence
XX of cancer in subject, by determining status of PHORI-All/PHORI-F5D6
XX gene products in tissue sample from subject and comparing it to normal
XX sample -
XX Claim 49; Page 195; 250pp; English.
XX CC The present invention relates to the isolation of novel human genes
XX designated PHORI-All and PHORI-F5D6 and their encoded proteins. The
XX gene encoding PHORI-All maps to chromosome 1q23, and the gene encoding
XX PHORI-F5D6 maps to chromosome 7q33-q35. The PHORI-All and PHORI-F5D6
XX polynucleotide and polypeptide sequences are useful in diagnostic and
XX therapeutic methods, and compositions for various cancers such as
XX prostate cancer. The sequences are useful for inhibiting the growth of
XX cancer cells that express PHORI-All or PHORI-F5D6 and for treating
XX cancer. The PHORI-All or PHORI-F5D6 polypeptide or a fragment thereof
XX can be used to elicit an immune response. AAU91563-AAU92962 represent
XX PHORI-All or PHORI-F5D6 peptides tested for HLA binding.

XX Sequence 9 AA;

Query Match 30.6%; Score 30; DB 23; Length 9;
Best Local Similarity 57.1%; Pred. No. 7.8e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIOY 13
||| :||
Db 2 ICHPLRY 8

RESULT 10
AAU92875
ID AAU92875 standard; peptide; 9 AA.

XX AC AAU92875;
XX DT 02-JUL-2002 (first entry)
XX DE PHORI-F5D6 peptide #613 tested for HLA binding.
XX KW Human; PHORI-All; PHORI-F5D6; prostate cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO200214501-A2.
XX PD 21-FEB-2002.
XX PF 17-AUG-2001; 2001WO-US25862.

XX 17-AUG-2000; 2000US-226241P.
PR (AGEN-) AGENSYS INC.
PA Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
PI Jakobovits A;
PI WPI; 2002-269193/31.
XX Monitoring PHOR1-A11/PHOR1-F5D6 gene products for monitoring presence
PT of cancer in subject, by determining status of PHOR1-A11/PHOR1-F5D6
PT gene products in tissue sample from subject and comparing it to normal
PT sample
XX Claim 49; Page 199; 250pp; English.
PS The present invention relates to the isolation of novel human genes
XX designated PHOR1-A11 and PHOR1-F5D6 and their encoded proteins. The
CC gene encoding PHOR1-A11 maps to chromosome 1q23, and the gene encoding
CC PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-A11 and PHOR1-F5D6
CC polynucleotide and polypeptide sequences are useful in diagnostic and
CC therapeutic methods, and compositions for various cancers such as
CC prostate cancer. The sequences are useful for inhibiting the growth of
CC cancer cells that express PHOR1-A11 or PHOR1-F5D6 and for treating
CC cancer. The PHOR1-A11 or PHOR1-F5D6 polypeptide or a fragment thereof
CC can be used to elicit an immune response. AAU91563-AAU92962 represent
CC PHOR1-A11 or PHOR1-F5D6 peptides tested for HLA binding.
XX Sequence 9 AA;
SQ Query Match 30.6%; Score 30; DB 23; Length 9;
Best Local Similarity 57.1%; Pred. No. 7.8e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 7 ICHQIQY 13
Db 3 ICHPLRY 9
RESULT 11
AAU92318
ID AAU92318 standard; peptide; 10 AA.
XX AC AAU92318;
XX DT 02-JUL-2002 (first entry)
XX DE PHOR1-F5D6 peptide #56 tested for HLA binding.
XX KW Human; PHOR1-A11; PHOR1-F5D6; prostate cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO200214501-A2.
XX PD 21-FEB-2002.
XX PF 17-AUG-2001; 2001WO-US25862.
XX PR 17-AUG-2000; 2000US-226241P.
XX PA (AGEN-) AGENSYS INC.
XX PI Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
XX PI Jakobovits A;
XX OS Homo sapiens.
XX PN WO200214501-A2.
XX PD 21-FEB-2002.
XX PF 17-AUG-2001; 2001WO-US25862.
XX PR 17-AUG-2000; 2000US-226241P.
XX PA (AGEN-) AGENSYS INC.
XX PI Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
XX PI Jakobovits A;
XX WPI; 2002-269193/31.
XX Monitoring PHOR1-A11/PHOR1-F5D6 gene products for monitoring presence
PT of cancer in subject, by determining status of PHOR1-A11/PHOR1-F5D6
PT gene products in tissue sample from subject and comparing it to normal
PT sample

XX Claim 49; Page 177; 250pp; English.
PS The present invention relates to the isolation of novel human genes
XX designated PHOR1-A11 and PHOR1-F5D6 and their encoded proteins. The
CC gene encoding PHOR1-A11 maps to chromosome 1q23, and the gene encoding
CC PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-A11 and PHOR1-F5D6
CC polynucleotide and polypeptide sequences are useful in diagnostic and
CC therapeutic methods, and compositions for various cancers such as
CC prostate cancer. The sequences are useful for inhibiting the growth of
CC cancer cells that express PHOR1-A11 or PHOR1-F5D6 and for treating
CC cancer. The PHOR1-A11 or PHOR1-F5D6 polypeptide or a fragment thereof
CC can be used to elicit an immune response. AAU91563-AAU92962 represent
CC PHOR1-A11 or PHOR1-F5D6 peptides tested for HLA binding.
XX Sequence 10 AA;
SQ Query Match 30.6%; Score 30; DB 23; Length 10;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 7 ICHQIQY 13
Db 4 ICHPLRY 10
RESULT 12
AAU92332
ID AAU92332 standard; peptide; 10 AA.
XX AC AAU92332;
XX DT 02-JUL-2002 (first entry)
XX DE PHOR1-F5D6 peptide #70 tested for HLA binding.
XX KW Human; PHOR1-A11; PHOR1-F5D6; prostate cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO200214501-A2.
XX PD 21-FEB-2002.
XX PF 17-AUG-2001; 2001WO-US25862.
XX PR 17-AUG-2000; 2000US-226241P.
XX PA (AGEN-) AGENSYS INC.
XX PI Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
XX PI Jakobovits A;
XX WPI; 2002-269193/31.
XX Monitoring PHOR1-A11/PHOR1-F5D6 gene products for monitoring presence
PT of cancer in subject, by determining status of PHOR1-A11/PHOR1-F5D6
PT gene products in tissue sample from subject and comparing it to normal
PT sample
XX Claim 49; Page 177; 250pp; English.
PS The present invention relates to the isolation of novel human genes
XX designated PHOR1-A11 and PHOR1-F5D6 and their encoded proteins. The
CC gene encoding PHOR1-A11 maps to chromosome 1q23, and the gene encoding
CC PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-A11 and PHOR1-F5D6
CC polynucleotide and polypeptide sequences are useful in diagnostic and
CC therapeutic methods, and compositions for various cancers such as
CC prostate cancer. The sequences are useful for inhibiting the growth of
CC cancer cells that express PHOR1-A11 or PHOR1-F5D6 and for treating
CC cancer. The PHOR1-A11 or PHOR1-F5D6 polypeptide or a fragment thereof
CC can be used to elicit an immune response. AAU91563-AAU92962 represent
CC PHOR1-A11 or PHOR1-F5D6 peptides tested for HLA binding.

XX Sequence 10 AA;

Query Match 30.6%; Score 30; DB 23; Length 10;
Best Local Similarity 57.1%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 7 ICHQIQY 13
||| :||
Db 1 ICHPLRY 7

RESULT 13

AAU92548
ID AAU92548 standard; peptide; 10 AA.

XX AC AAU92548;
XX DT 02-JUL-2002 (first entry)
XX DE PHOR1-F5D6 peptide #286 tested for HLA binding.
XX KW Human; PHOR1-All; PHOR1-F5D6; prostate cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO200214501-A2.
XX PD 21-FEB-2002.
XX PF 17-AUG-2001; 2001WO-US25862.
XX PR 17-AUG-2000; 2000US-226241P.
XX PA (AGEN-) AGENSYS INC.
XX PI Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
XX PJ Jakobovits A;
XX DR WPI; 2002-269193/31.
XX PT Monitoring PHOR1-All/PHOR1-F5D6 gene products for monitoring presence
XX of cancer in subject, by determining status of PHOR1-All/PHOR1-F5D6
XX gene products in tissue sample from subject and comparing it to normal
XX sample -
XX PS Claim 49; Page 186; 250pp; English.

XX The present invention relates to the isolation of novel human genes
XX designated PHOR1-All and PHOR1-F5D6 and their encoded proteins. The
XX gene encoding PHOR1-All maps to chromosome 1q23, and the gene encoding
XX PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-All and PHOR1-F5D6
XX polynucleotide and polypeptide sequences are useful in diagnostic and
XX therapeutic methods, and compositions for various cancers such as
XX prostate cancer. The sequences are useful for inhibiting the growth of
XX cancer cells that express PHOR1-All or PHOR1-F5D6 and for treating
XX cancer. The PHOR1-All or PHOR1-F5D6 polypeptide or a fragment thereof
XX can be used to elicit an immune response. AAU91563-AAU92962 represent
XX PHOR1-All or PHOR1-F5D6 peptides tested for HLA binding.

XX Query Match 30.6%; Score 30; DB 23; Length 10;
XX Best Local Similarity 57.1%; Pred. No. 2.4e+02;
XX Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 7 ICHQIQY 13
||| :||
Db 4 ICHPLRY 10

RESULT 14

AAU92639
ID AAU92639 standard; peptide; 10 AA.

XX AC AAU92639;
XX DT 02-JUL-2002 (first entry)
XX DE PHOR1-F5D6 peptide #377 tested for HLA binding.
XX KW Human; PHOR1-All; PHOR1-F5D6; prostate cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO200214501-A2.
XX PD 21-FEB-2002.
XX PF 17-AUG-2001; 2001WO-US25862.
XX PR 17-AUG-2000; 2000US-226241P.
XX PA (AGEN-) AGENSYS INC.
XX PI Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
XX PJ Jakobovits A;
XX DR WPI; 2002-269193/31.
XX PT Monitoring PHOR1-All/PHOR1-F5D6 gene products for monitoring presence
XX of cancer in subject, by determining status of PHOR1-All/PHOR1-F5D6
XX gene products in tissue sample from subject and comparing it to normal
XX sample -
XX PS Claim 49; Page 186; 250pp; English.

XX The present invention relates to the isolation of novel human genes
XX designated PHOR1-All and PHOR1-F5D6 and their encoded proteins. The
XX gene encoding PHOR1-All maps to chromosome 1q23, and the gene encoding
XX PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-All and PHOR1-F5D6
XX polynucleotide and polypeptide sequences are useful in diagnostic and
XX therapeutic methods, and compositions for various cancers such as
XX prostate cancer. The sequences are useful for inhibiting the growth of
XX cancer cells that express PHOR1-All or PHOR1-F5D6 and for treating
XX cancer. The PHOR1-All or PHOR1-F5D6 polypeptide or a fragment thereof
XX can be used to elicit an immune response. AAU91563-AAU92962 represent
XX PHOR1-All or PHOR1-F5D6 peptides tested for HLA binding.

XX Query Match 30.6%; Score 30; DB 23; Length 10;
XX Best Local Similarity 57.1%; Pred. No. 2.4e+02;
XX Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 7 ICHQIQY 13
||| :||
Db 4 ICHPLRY 10

XX Query Match 30.6%; Score 30; DB 23; Length 10;
XX Best Local Similarity 57.1%; Pred. No. 2.4e+02;
XX Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 7 ICHQIQY 13
||| :||
Db 4 ICHPLRY 10

RESULT 14
AAU92639

ID AAU92639 standard; peptide; 10 AA.

XX AC AAU92639;
XX DT 02-JUL-2002 (first entry)
XX DE PHOR1-F5D6 peptide #377 tested for HLA binding.
XX KW Human; PHOR1-All; PHOR1-F5D6; prostate cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO200214501-A2.
XX PD 21-FEB-2002.
XX PF 17-AUG-2001; 2001WO-US25862.
XX PR 17-AUG-2000; 2000US-226241P.
XX PA (AGEN-) AGENSYS INC.
XX PI Hubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
XX PJ Jakobovits A;
XX DR WPI; 2002-269193/31.
XX PT Monitoring PHOR1-All/PHOR1-F5D6 gene products for monitoring presence
XX of cancer in subject, by determining status of PHOR1-All/PHOR1-F5D6
XX gene products in tissue sample from subject and comparing it to normal
XX sample -
XX PS Claim 49; Page 189; 250pp; English.

XX The present invention relates to the isolation of novel human genes
XX designated PHOR1-All and PHOR1-F5D6 and their encoded proteins. The
XX gene encoding PHOR1-All maps to chromosome 1q23, and the gene encoding
XX PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-All and PHOR1-F5D6
XX polynucleotide and polypeptide sequences are useful in diagnostic and
XX therapeutic methods, and compositions for various cancers such as
XX prostate cancer. The sequences are useful for inhibiting the growth of
XX cancer cells that express PHOR1-All or PHOR1-F5D6 and for treating
XX cancer. The PHOR1-All or PHOR1-F5D6 polypeptide or a fragment thereof
XX can be used to elicit an immune response. AAU91563-AAU92962 represent
XX PHOR1-All or PHOR1-F5D6 peptides tested for HLA binding.

XX Query Match 30.6%; Score 30; DB 23; Length 10;
XX Best Local Similarity 57.1%; Pred. No. 2.4e+02;
XX Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 7 ICHQIQY 13
||| :||
Db 4 ICHPLRY 10

RESULT 15

AAU92744
ID AAU92744 standard; peptide; 10 AA.

XX AC AAU92744;
XX DT 02-JUL-2002 (first entry)
XX DE PHOR1-F5D6 peptide #482 tested for HLA binding.
XX KW Human; PHOR1-All; PHOR1-F5D6; prostate cancer; cytostatic.
XX OS Homo sapiens.
XX PN WO200214501-A2.
XX

PD 21-FEB-2002.
 XX 17-AUG-2001; 2001WO-US25862.
 PF XX
 XX 17-AUG-2000; 2000US-226241P.
 PR XX
 XX (AGEN-) AGENSYS INC.
 PA
 XX Rubert RS, Raitano AB, Faris M, Challita-Eid PM, Ge W;
 PI Jakobovits A;
 XX WPI; 2002-269193/31.
 DR
 XX Monitoring PHOR1-All/PHOR1-F5D6 gene products for monitoring presence
 PT of cancer in subject, by determining status of PHOR1-All/PHOR1-F5D6
 PT gene products in tissue sample from subject and comparing it to normal
 PT sample -
 XX
 PS Claim 49; Page 193; 250pp; English.
 XX
 CC The present invention relates to the isolation of novel human genes
 CC designated PHOR1-All and PHOR1-F5D6 and their encoded proteins. The
 CC gene encoding PHOR1-All maps to chromosome 1q23, and the gene encoding
 CC PHOR1-F5D6 maps to chromosome 7q33-q35. The PHOR1-All and PHOR1-F5D6
 CC polynucleotide and polypeptide sequences are useful in diagnostic and
 CC therapeutic methods, and compositions for various cancers such as
 CC prostate cancer. The sequences are useful for inhibiting the growth of
 CC cancer cells that express PHOR1-All or PHOR1-F5D6 and for treating
 CC cancer. The PHOR1-All or PHOR1-F5D6 polypeptide or a fragment thereof
 CC can be used to elicit an immune response. AAU91563-AAU92962 represent
 CC PHOR1-All or PHOR1-F5D6 peptides tested for HLA binding.
 XX
 SQ Sequence 10 AA;
 Query Match 30.6%; Score 30; DB 23; Length 10;
 Best Local Similarity 57.1%; Pred. No. 2.4e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Qy 7 ICHOIQY 13
 Db 3 ICHPLRY 9

Search completed: April 23, 2003, 13:42:56
 Job time : 28.2022 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:21:27 ; Search time 27.9101 Seconds
(without alignments)
132.885 Million cell updates/sec

Title: us-09-836-073-2

Perfect score: 96
Sequence: 1 AAALQAICQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_proteus:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	86.5	381	11 Q9CYB9	Q9cyb9 mus musculus
2	70	72.9	286	13 Q8QH15	Q8qh15 gallus gall
3	52	54.2	390	5 Q8T8V5	Q8t8v5 drosophila
4	50	52.1	391	2 P72382	P72382 staphylococ
5	50	52.1	391	2 P95709	P95709 staphylococ
6	50	52.1	391	16 Q99X57	Q99x57 staphylococ
7	47	49.0	411	10 Q9FL36	Q9fl36 arabidopsis
8	47	49.0	422	10 Q9LA38	Q9la38 arabidopsis
9	45.5	47.4	928	10 Q9LJ02	Q9lj02 oryza sativ
10	45	46.9	396	5 O01806	O01806 caenorhabdi
11	45	46.9	688	5 Q9VBE6	Q9vbe6 drosophila
12	45	46.9	826	10 Q940X9	Q940x9 arabidopsis
13	44	45.8	389	10 Q8S0T8	Q8s0t8 oryza sativ
14	44	45.8	392	6 Q9GKK4	Q9gkk4 canis famil
15	44	45.8	506	10 Q9M143	Q9m143 arabidopsis
16	43	44.8	111	17 Q978R4	Q978r4 thermoplasm

17	43	44.8	119	10 Q9ZP89	Q9zp89 neurotheca
18	43	44.8	166	5 O44678	O44678 caenorhabdi
19	43	44.8	355	10 Q8RYB4	Q8ryb4 phytophthor
20	43	44.8	545	10 Q80567	Q80567 arabidopsis
21	43	44.8	569	10 Q49048	Q49048 arabidopsis
22	43	44.8	569	10 Q80650	Q80650 arabidopsis
23	42	43.8	159	16 Q97LL1	Q97ll1 clostridium
24	42	43.8	186	11 Q8VDM3	Q8vdm3 mus musculu
25	42	43.8	214	11 Q8CSI2	Q8csi2 mus musculu
26	42	43.8	239	3 Q9P8N6	Q9p8n6 cochlidiobol
27	42	43.8	328	4 Q9NW12	Q9nw12 homo sapien
28	42	43.8	342	11 Q9CU05	Q9cu05 mus musculu
29	42	43.8	482	16 Q8YX0	Q8yx0 bruceella me
30	42	43.8	669	16 Q8XRQ1	Q8xrq1 raistonia s
31	42	43.8	1379	5 Q9VAV5	Q9vav5 drosophila
32	42	43.8	1403	5 Q9NH6	Q9nh6 drosophila
33	41	42.7	132	16 Q926P7	Q926p7 listeria in
34	41	42.7	261	16 Q8XU67	Q8xu67 raistonia s
35	41	42.7	335	9 Q9XJR3	Q9xjr3 bacterioph
36	41	42.7	343	2 Q9XBI7	Q9xbi7 bacillus ce
37	41	42.7	469	16 Q66887	Q66887 aquifex aeo
38	41	42.7	505	10 Q9FQY6	Q9fqy6 capsicum an
39	41	42.7	523	10 Q94K80	Q94k80 arabidopsis
40	41	42.7	837	16 Q9PR25	Q9pr25 ureaplasma
41	41	42.7	1575	5 Q8SSN3	Q8ssn3 dictyostell
42	41	42.7	2160	5 Q17709	Q17709 caenorhabdi
43	40.5	42.2	453	5 Q9W438	Q9w438 drosophila
44	40	41.7	51	16 Q9K0R3	Q9k0r3 neisseria m
45	40	41.7	91	10 Q9AUF9	Q9auf9 brassica ca

ALIGNMENTS

RESULT 1

Q9CYB9 PRELIMINARY; PRT; 381 AA.
ID Q9CYB9
AC Q9CYB9
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Sjogren syndrome antigen B.
GN SSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavani T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schirni L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Sasaki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK017822; BAB30957.1;
DR MGD; MGI:98423; Ssb.

DR InterPro: IPR002344; Lupus.La.
 DR InterPro: IPR000504; RNA_rec_mot.
 DR Pfam: PF00076; rtm; 1.
 DR PRINTS: PR00302; LUPUSLA.
 SQ SEQUENCE 381 AA; 43951 MW; 2E2DEF1452C0F0E9 CRC64;

Query Match 86.5%; Score 83; DB 11; Length 381;
 Best Local Similarity 88.2%; Pred. No. 8.4e-06;
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 ALEAICQIQIYYFGDF 18
 |||||:|||||
 DB 12 ALEAKICQIQIYYFGDF 28

RESULT 2

ID Q8QHI5 PRELIMINARY; PRT; 206 AA.
 AC Q8QHI5;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Untranslated region binding-protein.
 GN UBP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA L'Ecuver T.J., Fang H.-L.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF467897; AAL76269.1;
 SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFFB90E9 CRC64;

Query Match 72.9%; Score 70; DB 13; Length 206;
 Best Local Similarity 80.0%; Pred. No. 0.00069;
 Matches 12; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 LEAICQIQIYYFGD 17
 |||||:|||||
 DB 13 LESKICQIQIYYFGN 27

RESULT 3

ID Q8T8V5 PRELIMINARY; PRT; 390 AA.
 AC Q8T8V5;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE AT22034p.
 GN LA.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY075257; AAL68124.1;
 SQ SEQUENCE 390 AA; 44869 MW; 797FDE26B903C909 CRC64;

Query Match 54.2%; Score 52; DB 5; Length 390;
 Best Local Similarity 64.3%; Pred. No. 1.4;

OY 4 EAQICQIQIYYFGD 17
 |||||:|||||
 DB 51 ERAIRQVEIYFGD 64

RESULT 4
 P72382 PRELIMINARY; PRT; 391 AA.
 AC P72382;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE CapSP.
 GN CAPSP.

OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BECKER;
 RX MEDLINE-96178981; PubMed-8606192;
 RA Sau S., Lee C.Y.;
 RT "Cloning of type 8 capsule genes and analysis of gene clusters for the
 RT production of different capsular polysaccharides in Staphylococcus
 RT aureus.";
 RL J. Bacteriol. 178:2118-2126(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BECKER;
 RX MEDLINE-97197525;
 RA Sau S., Sun J., Lee C.Y.;
 RT "Molecular characterization and transcriptional analysis of type 8
 RT capsule genes in Staphylococcus aureus.";
 RL J. Bacteriol. 179:1614-1621(1997).
 DR EMBL: U73374; AAB49445.1;
 DR HSP; P27828; IF6D.
 DR InterPro: IPR003331; Epimerase_2.
 DR Pfam: PF02350; Epimerase_2; 1.
 DR TIGRFAMs: TIGR00236; wecB; 1.
 SQ SEQUENCE 391 AA; 44247 MW; 1E8D9FAA9BC76F0D CRC64;

Query Match 52.1%; Score 50; DB 2; Length 391;
 Best Local Similarity 72.7%; Pred. No. 3.1;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 6 QICQIQIYYFG 16
 |||||:|||||
 DB 366 RICEAIEYFG 376

RESULT 5

P95709 PRELIMINARY; PRT; 391 AA.
 ID P95709;
 AC P95709;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE CapSP.
 GN CAPSP.
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-REYNOLDS, AND NEWMAN;
 RX MEDLINE-97388587; PubMed-9245821;
 RA Sau S., Bhasin N., Wann E.R., Lee J.C., Foster T.J., Lee C.Y.;
 RT "The Staphylococcus aureus allelic genetic loci for serotype 5 and 8
 RT capsule expression contain the type-specific genes flanked by common


```
RT Genes.;
RL Microbiology 143:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-REYNOLDS, AND NEWMAN;
RX MEDLINE-98101481; PubMed-9440531;
RA Kiser K.B., Lee J.C.;
RT "Staphylococcus aureus cap50 and cap5p genes functionally complement
RT mutations affecting enterobacterial common-antigen biosynthesis in
RT Escherichia coli.";
RL J. Bacteriol. 180:403-406(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-REYNOLDS, AND NEWMAN;
RX MEDLINE-98125727; PubMed-9466251;
RA Bhasin N., Albus A., Michon F., Livolsi P.J., Park J.-S., Lee J.C.;
RT "Identification of a gene essential for O-acetylation of the
RT Staphylococcus aureus type 5 capsular polysaccharide.";
RL Mol. Microbiol. 27:9-21(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-REYNOLDS, AND NEWMAN;
RA Bagga N., Wann E.R., Foster T.J., Lee J.C.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U81973; AAC46099.1; -
DR HSSP; P27828; 1F6D.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMs; TIGR00236; wecB; 1.
SQ SEQUENCE 391 AA; 44328 MW; B0105F690DB7CF1D CRC64;

Query Match 52.1%; Score 50; DB 2; Length 391;
Best Local Similarity 72.7%; Pred. No. 3.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 QICQOIEYFYG 16
DB 366 RICEAIEYFYG 376

RESULT 6
Q99X57 PRELIMINARY; PRT; 391 AA.
AC Q99X57;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Capsular polysaccharide synthesis enzyme Cap5p.
GN CAP OR SAV0164 OR SA0159.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OC Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE-21311952; PubMed-1141846;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003358; BAB56326.1; -
DR EMBL; AF003129; BAB41379.1; -
DR HSSP; P27828; 1F6D.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.

DR TIGRFAMs; TIGR00236; wecB; 1.
KW Complete proteome.
SQ SEQUENCE 391 AA; 44372 MW; D0DF5FA715BCC6CC CRC64;

Query Match 52.1%; Score 50; DB 16; Length 391;
Best Local Similarity 72.7%; Pred. No. 3.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 QICQOIEYFYG 16
DB 366 RICEAIEYFYG 376

RESULT 7
Q9FL36 PRELIMINARY; PRT; 411 AA.
AC Q9FL36;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similarity to RNA-binding protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE-98344145; PubMed-9679202;
RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT features of the regions of 1,381,565 bp covered by twenty one
RT physically assigned pi and TAC clones.";
RL DNA Res. 5:131-145(1998).
DR EMBL; AB010698; BAB11080.1; -
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00360; RRM; 1.
DR POSITE; PSS0102; RRM; 1.
DR SMART; PSS0102; RRM; 1.
SQ SEQUENCE 411 AA; 45655 MW; A2EF62EB5B9B099 CRC64;

Query Match 49.0%; Score 47; DB 10; Length 411;
Best Local Similarity 53.3%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LEAQICQOIEYFGD 17
DB 92 LNQKIRQVEYFSD 106

RESULT 8
Q94A38 PRELIMINARY; PRT; 422 AA.
AC Q94A38;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE A5946250/MEP112.3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
```

```

RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY050403; AK91419.1;
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 1.
DR PROSITE: PS0102; RRM; 1.
SQ SEQUENCE 422 AA; 46842 MW; 4EC4BBF1E068F0E CRC64;

Query Match 49.0%; Score 47; DB 10; Length 422;
Best Local Similarity 53.3%; Pred. No. 11;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LEAQICQIEYFGD 17
DB 103 LNQIIRQVEYFGD 117
I : : : : :
: : : : :

RESULT 9
Q9LJ02 PRELIMINARY; PRT; 928 AA.
AC Q9LJ02;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Similar to KIAA0731 protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:PO499C11."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP001080; BAA90356.1;
SQ SEQUENCE 928 AA; 103745 MW; 53FC46E24A446EB4 CRC64;

Query Match 47.4%; Score 45.5; DB 10; Length 928;
Best Local Similarity 62.5%; Pred. No. 43;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 3 LEAQICQIEYFGD 17
DB 278 LRKILQVEYFGD 293
I : : : : :
: : : : :

RESULT 10
O01806 PRELIMINARY; PRT; 396 AA.
AC O01806;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE C44E4.4 protein.
GN C44E4.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA None;
RX MEDLINE=99069613; PubMed=9851916;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."

```

```

RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Sammons L., Wohlmann P., Gillam B.;
RT "The sequence of C. elegans cosmid C4E4."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003140; AAB54169.1;
DR InterPro: IPR002344; Lupus_Ia.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 1.
DR PRINTS: PR00302; LUPUSLA.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS0102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
SQ SEQUENCE 396 AA; 43631 MW; A2D828A4FAA3C34 CRC64;

Query Match 46.9%; Score 45; DB 5; Length 396;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 4 EAQICQIEYFGD 17
DB 13 DQIKIQEYFGN 26
I : : : : :
: : : : :

RESULT 11
Q9VBE6 PRELIMINARY; PRT; 688 AA.
AC Q9VBE6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CS4555 protein (GHI5286P).
GN CS455.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balwle R.M., Basu A., Baxendale J., Bayraktarglu L., Beasley E.M.,
RA Beeson K.F., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

```

Query Match	46.9%;	Score 45;	DB 10;	Length 826;
Best Local Similarity	46.7%;	Pred. No. 47;		

```

RESULT 14
Q9GKK4
ID Q9GKK4 PRELIMINARY; PRT; 392 AA.
AC Q9GKK4;
DT 01-WAR-2001 (TrenBLrel. 16, Created)
DT 01-WAR-2001 (TrenBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrenBLrel. 21, Last annotation update)
DE Galactokinase (EC 2.7.1.6).
GN GALK1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
[1]
RN SEQUENCE FROM N.A.
RT Identification and analysis of canine galactokinase (GALK1) cdna." ;
RA "IdentJan D.J." ;
RL Cited (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + D-GALACTOSE -> ADP + D-GALACTOSE 1-
CC PHOSPHATE.
CC -1- PATHWAY: FIRST REACTION OF GALACTOSE METABOLISM.
CC -1- SIMILARITY: BELONGS TO THE GMP KINASE FAMILY.
CC EMBL; AF213513; AAG43832.1; -.
DR InterPro; IPR00705; Galactokinase.
DR InterPro; IPR001174; Galkinase.
DR InterPro; IPR001745; GHMPKase.ATP.
DR InterPro; IPR001459; Mv_gal_kin.
DR Pfam; PF00288; GHMP_kinases; 1.
DR PRINTS; PR00473; GALCTOKINASE.
DR PRINTS; PR00960; LMBPPROTEIN.
DR PRINTS; PR00959; MEYGALKINASE.
DR TIGRFS; TIGR00131; gal_kin; 1.
DR PROSITE; PS00106; GALACTOKINASE; 1.

```

DR PROSITE: PS00627; GHMP_KINASES_ATP; 1.
KW ATP-binding, Galactose metabolism, Kinase, Transferase.
SQ SEQUENCE 392 AA; 42120 MW; 8FD5ECF3AB4386BF CRC64;

Query Match 45.8%; Score 44; DB 6; Length 392;
Best Local Similarity 57.1%; Pred. No. 32;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ALEAQICQIEYVF 15
| | | | | | | | | |
DB 164 ARAQVQQAHSF 177

RESULT 15

Q9M143 PRELIMINARY; PRT; 506 AA.
AC Q9M143;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE Putative RING zinc finger protein.
GN AR4G01270.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]

RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL: AL161491; CAB80936.1;
DR InterPro: IPR001841; Znf_finger.
DR Pfam: PF00097; zf-C3HC4; 1.
DR SMART: SM00184; RING; 1.
KW Zinc-finger.
SQ SEQUENCE 506 AA; 56518 MW; 1B06751A4E299FE8 CRC64;
Query Match 45.8%; Score 44; DB 10; Length 506;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 AQICQIEYVF 16
| | | | | | | | | |
DB 491 SQGCLQIEHYFG 502

Search completed: April 23, 2003, 13:32:47
Job time : 29.9101 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: April 23, 2003, 13:27:28 ; Search time 10.3146 Seconds
(without alignments)
139.834 Million cell updates/sec

Title: US-09-836-073-2

Perfect score: 96

Sequence: 1 AALEAQICQIEYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/ECTUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	18	9	US-09-836-073-2
2	87	90.6	18	9	US-09-836-073-1
3	87	90.6	18	9	US-09-836-073-14
4	87	90.6	460	9	US-10-102-806-695
5	83	86.5	17	9	US-09-836-073-13
6	81	84.4	18	9	US-09-836-073-9
7	79	82.3	18	9	US-09-836-073-11
8	79	82.3	18	9	US-09-836-073-12
9	78	81.2	19	9	US-09-836-073-10
10	76.5	79.7	19	9	US-09-836-073-16
11	75	78.1	18	9	US-09-836-073-3
12	75	78.1	18	9	US-09-836-073-4
13	74	77.1	18	9	US-09-836-073-15
14	72	75.0	18	9	US-09-836-073-7
15	71	74.0	18	9	US-09-836-073-8
16	63	65.6	18	9	US-09-836-073-5
17	61	63.5	38	9	US-09-843-676-25
18	61	63.5	38	9	US-09-766-253-25
19	61	63.5	38	9	US-09-438-486-25

20	61	63.5	38	9	US-10-053-758-25	Sequence 25, Appl
21	61	63.5	38	9	US-10-054-295-25	Sequence 25, Appl
22	61	63.5	38	9	US-10-054-611-25	Sequence 25, Appl
23	58	60.4	16	9	US-09-836-073-19	Sequence 19, Appl
24	56	58.3	18	9	US-09-836-073-6	Sequence 6, Appl
25	50.5	52.6	37	9	US-09-843-676-24	Sequence 24, Appl
26	50.5	52.6	37	9	US-09-766-253-24	Sequence 24, Appl
27	50.5	52.6	37	9	US-09-438-486-24	Sequence 24, Appl
28	50.5	52.6	37	9	US-10-053-758-24	Sequence 24, Appl
29	50.5	52.6	37	9	US-10-054-295-24	Sequence 24, Appl
30	50.5	52.6	37	9	US-10-054-611-24	Sequence 24, Appl
31	49	51.0	39	9	US-09-843-676-26	Sequence 26, Appl
32	49	51.0	39	9	US-09-766-253-26	Sequence 26, Appl
33	49	51.0	39	9	US-09-438-486-26	Sequence 26, Appl
34	49	51.0	39	9	US-10-053-758-26	Sequence 26, Appl
35	49	51.0	39	9	US-10-054-295-26	Sequence 26, Appl
36	49	51.0	39	9	US-10-054-611-26	Sequence 26, Appl
37	45	46.9	18	9	US-09-836-073-17	Sequence 17, Appl
38	44	45.8	406	9	US-10-102-806-668	Sequence 668, App
39	43.5	45.3	18	9	US-09-836-073-18	Sequence 18, Appl
40	42	43.8	211	10	US-09-925-302-786	Sequence 786, App
41	42	43.8	487	9	US-10-023-437-19	Sequence 19, Appl
42	41	42.7	229	10	US-09-893-737-318	Sequence 318, App
43	41	42.7	382	10	US-09-893-737-36	Sequence 36, Appl
44	40	41.7	262	9	US-09-813-453A-8	Sequence 8, Appl
45	40	41.7	505	10	US-09-765-873A-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; LENGTH: 18
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match 100.0%; Score 96; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAQICQIEYFGDF 18

Db 1 AALEAQICQIEYFGDF 18

RESULT 2

US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-1

Query Match 90.6%; Score 87; DB 9; Length 18;
 Best Local Similarity 88.9%; Pred. No. 3.3e-07;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQICQIEYYFGDF 18
 Db 1 AALEAKICHQIEYYFGDF 18

RESULT 3
 US-09-836-073-14
 ; Sequence 14, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 22002034822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Bovine
 US-09-836-073-14

Query Match 90.6%; Score 87; DB 9; Length 18;
 Best Local Similarity 88.9%; Pred. No. 3.3e-07;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQICQIEYYFGDF 18
 Db 1 AALEAKICHQIEYYFGDF 18

RESULT 4
 US-10-102-806-695
 ; Sequence 695, Application US/10102806
 ; Publication No. US20030054421A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: P103P1C1
 ; CURRENT APPLICATION NUMBER: US/10/102,806
 ; CURRENT FILING DATE: 2002-03-22
 ; PRIOR APPLICATION NUMBER: 09/925,298
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05881
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 846
 ; SOFTWARE: Patent Ver. 2.0
 ; SEQ ID NO 695
 ; LENGTH: 460
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

US-10-102-806-695

Query Match 90.6%; Score 87; DB 9; Length 460;
 Best Local Similarity 88.9%; Pred. No. 6.9e-06;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQICQIEYYFGDF 18
 Db 63 AALEAKICHQIEYYFGDF 80

RESULT 5
 US-09-836-073-13
 ; Sequence 13, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 22002034822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 17
 ; TYPE: PRT
 ; ORGANISM: Mouse
 US-09-836-073-13

Query Match 86.5%; Score 83; DB 9; Length 17;
 Best Local Similarity 88.2%; Pred. No. 1.3e-06;
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AALEAQICQIEYYFGDF 18
 Db 1 AALEAKICHQIEYYFGDF 17

RESULT 6
 US-09-836-073-9
 ; Sequence 9, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baidya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 22002034822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-9

Query Match 84.4%; Score 81; DB 9; Length 18;
 Best Local Similarity 88.2%; Pred. No. 2.8e-06;
 Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQICQIEYYFGD 17
 Db 1 AALEAKICHQIEYYFGD 17

```
RESULT 7
US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match      82.3%; Score 79; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 5.7e-06;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 AALEAQICQIQIYYFGDF 18
        |||||:|||||
Db      1 AALEAKICHQIEYQGF 18

RESULT 8
US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match      82.3%; Score 79; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 5.7e-06;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 AALEAQICQIQIYYFGDF 18
        |||||:|||||
Db      1 AALEAKICHQIEYQGF 18

RESULT 9
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match      79.7%; Score 76.5; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 1.5e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY      1 AALEAQICQIQIYYFGDF 18
        |||||:|||||
Db      1 AALEAKICHQIEYQGF 19

RESULT 11
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3
```

Query Match 78.1%; Score 75; DB 9; Length 18;
Best Local Similarity 72.2%; Pred. No. 2.4e-05;
Matches 13; Conservative 3; Mismatches 2; Indels

```

QY      1 AALEAQCQI EYFGDF 18
      |||:| | | | | |
Db      1 AALQAKICH OIQYFGQF 18

```

RESULT 12
US-09-836-073-4
: Sequence 4, Application US/09836073

```

; Patent No. US20020173475A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Dasgupta, Asim
;
; APPLICANT: Das, S.
;
; APPLICANT: Baidya, Narayan
;
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
;
; FILE REFERENCE: 22002054822
;
; CURRENT APPLICATION NUMBER: US/09/836,073
;
; CURRENT FILING DATE: 2002-10-24
;
; PRIOR APPLICATION NUMBER: 09/316,630
;
; PRIOR FILING DATE: 1999-05-21
;
; NUMBER OF SEQ ID NOS: 19
;
; SOFTWARE: FastSeq for Windows Version 4.0
;

```

```

; SEQ ID NO. 4
; LENGTH: 18

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-09-836-073-4

Query Match 78.1%; Score 75; DB 9; Length 18;
Best Local Similarity 86.7%; Pred. NO. 2.4e-05;
Matches 13; Conservative 1; Mismatches 1; Indels

QY 4 EAQICQIEYFEGDF 18
 11:11 11111111
 Db 4 EAKICHOIEYFEGDF 18

RESULT 13
US-09-836-073-15
: Sequence 15. Application US/09836073

```

Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Baidya, Narayan
TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
FILE REFERENCE: 220002054822
CURRENT APPLICATION NUMBER: US/09/836,073
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/316,630
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 19
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15

```

```

; LENGTH: 18
; TYPE: PRT

```

; ORGANISM: Xenopus
US-09-836-073-15

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Query Match 77.1%; Score 74; DB 9; Length 18;
Best Local Similarity 75.0%; Pred. NO. 3.4e-05;
Matches 12; Conservative 3; Mismatches 1; Indels

QY 3 LEAQICQIEYYFGDF 18
| : : | : | : | : | : |
Db 3 LDTKICEOIEVVEGDF 18

RESULT 14
US-09-936-073-7

; Sequence 7, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:

```

1 APPLICANT: Dasgupta, Asim
2 APPLICANT: Das, S.
3 APPLICANT: Baigya, Narayan
4 TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
5 FILE REFERENCE: 220002054822
6 CURRENT APPLICATION NUMBER: US/09/835,073
7 CURRENT FILING DATE: 2002-10-24
8 PRIOR APPLICATION NUMBER: 09/316,630
9 PRIOR FILING DATE: 1999-05-21
10 NUMBER OF SEQ ID NOS: 19
11 SOFTWARE: FastSeq for Windows Version 4.0
12 SEQ ID NO 7
13 LENGTH: 18

```

;
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-7

Query Match 75.0%; Score 72; DB 9; Length 18;
Best Local Similarity 82.4%; Pred. No. 6.9e-05;
Matches 14: Conservative 1; Mismatches 2; Indels

```

QY      1 AALEAQICQIEYYFGD 17
        |||||:|||||
db      1 AALEAKICHOIEYYOGD 17

```

RESULT 15

US-09-836-073-8
; Sequence 8, Application US/09836073
; Patent No. US20020173475A1

```

; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baldya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8

```

```

; LENGTH: 18
; TYPE: PRT

```

ORGANISM: Homo Sapiens
US-09-836-073-8

Query Match 74.0%; Score 71; DB 9; Length 18;
Best Local Similarity 77.8%; Pred. No. 9.8e-05;
Matches 14; Conservative 1; Mismatches 3; Indels

```
QY      1 AALEAQCQIEYYFGDF 18
          |||||:|||||
Db      1 AALEAKICHOLEOOFGDF 18
```

Search completed: April 23, 2003, 13:38:19
Job time : 11.3146 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:24 ; Search time 11.2247 Seconds
(without alignments)
47.183 Million cell updates/sec

Title: US-09-836-073-2
Perfect score: 96
Sequence: 1 AALEAQICQIIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/BCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	87	90.6	18	4	US-09-316-630-3	Sequence 3, Appli
2	87	90.6	18	4	US-09-316-630-4	Sequence 4, Appli
3	61	63.5	38	3	US-08-851-843A-25	Sequence 25, Appl
4	61	63.5	38	4	US-08-974-549A-215	Sequence 215, App
5	61	63.5	38	4	US-08-854-050-25	Sequence 25, Appl
6	61	63.5	38	4	US-09-430-323-25	Sequence 25, Appl
7	60	62.5	38	4	US-08-974-549A-214	Sequence 214, App
8	50.5	52.6	37	3	US-08-851-843A-24	Sequence 24, Appl
9	50.5	52.6	37	4	US-08-854-050-24	Sequence 24, Appl
10	50.5	52.6	37	4	US-09-430-323-24	Sequence 24, Appl
11	49	51.0	39	3	US-08-851-843A-26	Sequence 26, Appl
12	49	51.0	39	4	US-08-974-549A-216	Sequence 216, App
13	49	51.0	39	4	US-08-854-050-26	Sequence 26, Appl
14	49	51.0	39	4	US-09-430-323-26	Sequence 26, Appl
15	44	45.8	392	1	US-08-451-777A-33	Sequence 33, Appl
16	44	45.8	392	2	US-08-451-778A-33	Sequence 33, Appl
17	44	45.8	392	2	US-08-998-208-33	Sequence 33, Appl
18	41	42.7	29	1	US-08-451-777A-3	Sequence 3, Appli
19	41	42.7	29	2	US-08-451-778A-3	Sequence 3, Appli
20	41	42.7	29	2	US-08-998-208-3	Sequence 3, Appli
21	41	42.7	29	5	PCT-US94-10825-3	Sequence 3, Appli
22	41	42.7	29	5	PCT-US95-06743-3	Sequence 3, Appli
23	40	41.7	505	4	US-09-627-216A-12	Sequence 12, Appl
24	40	41.7	505	4	US-09-126-420A-22	Sequence 22, Appl
25	40	41.7	754	2	US-08-941-262-1	Sequence 1, Appli
26	40	41.7	754	2	US-08-941-262-3	Sequence 3, Appli
27	40	41.7	1180	3	US-09-224-024-28	Sequence 28, Appl

28	40	41.7	1180	5	PCT-US94-07902-28	Sequence 28, Appl
29	39	40.6	456	1	US-08-205-719-4	Sequence 4, Appli
30	39	40.6	456	4	US-08-431-517F-6	Sequence 6, Appli
31	39	40.6	482	4	US-08-431-517F-5	Sequence 5, Appli
32	39	40.6	482	6	5245013-2	Patent No. 5245013
33	38.5	40.1	407	1	US-08-487-823B-2	Sequence 2, Appli
34	38.5	40.1	407	2	US-08-997-040-2	Sequence 2, Appli
35	38.5	40.1	407	2	US-09-203-237-2	Sequence 2, Appli
36	38.5	40.1	410	3	US-08-948-997-2	Sequence 2, Appli
37	38.5	40.1	410	4	US-09-348-817A-2	Sequence 2, Appli
38	38.5	40.1	410	4	US-09-348-817A-3	Sequence 3, Appli
39	38	39.6	38	3	US-08-851-843A-27	Sequence 27, App
40	38	39.6	38	4	US-08-974-549A-217	Sequence 217, App
41	38	39.6	38	4	US-08-854-050-27	Sequence 27, Appl
42	38	39.6	38	4	US-09-430-323-27	Sequence 27, Appl
43	38	39.6	379	4	US-09-603-185-6	Sequence 6, Appli
44	38	39.6	556	2	US-08-505-377-1	Sequence 1, Appli
45	38	39.6	556	3	US-08-798-269-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match 90.6%; Score 87; DB 4; Length 18;
Best Local Similarity 88.9%; Pred.No. 2.7e-08;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AALEAQICQIIEYFGDF 18
|||||:|||||
Db 1 AALEAKICHQIIEYFGDF 18

RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

;; PRIOR FILING DATE: 1997-10-06
;; PRIOR APPLICATION NUMBER: 08/321,427
;; PRIOR FILING DATE: 1994-10-11
;; PRIOR APPLICATION NUMBER: 60/086,527
;; PRIOR FILING DATE: 1998-05-22
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 18
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
;; OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 90.6%; Score 87; DB 4; Length 18;
Best Local Similarity 88.9%; Pred. No. 2.7e-08;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIQIYYFGDF 18
|||:|||||
Db 1 AALEAKICQIYYFGDF 18

RESULT 3
US-08-851-843A-25
; Sequence 25, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200

;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
US-08-851-843A-25

Query Match 63.5%; Score 61; DB 3; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0013;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICQIQIYYFGD 17
||:|||||
Db 1 ICEQIYYFGD 11

RESULT 4
US-08-974-549A-215
; Sequence 215, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US97/17618
;; FILING DATE: 01-OCT-1997
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/US97/17885
;; FILING DATE: 01-OCT-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002610US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 215:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-974-549A-215

Query Match 63.5%; Score 61; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0013;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICQIEYYFGD 17
II:IIIIIIII
Db 1 ICEQIEYYFGD 11

RESULT 5
US-08-854-050-25
; Sequence 25, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-002930US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 38 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
US-08-854-050-25

Query Match 63.5%; Score 61; DB 4; Length 38;
Best Local Similarity 90.9%; Pred. No. 0.0013;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICQIEYYFGD 17
II:IIIIIIII
Db 1 ICEQIEYYFGD 11

RESULT 6
US-09-430-323-25
; Sequence 25, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-0029300S

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 38 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 25:

US-09-430-323-25

Query Match 63.5%; Score 61; DB 4; Length 38;

Best Local Similarity 90.9%; Pred. No. 0.0013;

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICQQEYFYFGD 17

Db 1 ICEQIEYFGD 11

RESULT 7

US-08-974-549A-214

Sequence 214, Application US/08974549A

Patent No. 6166178

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin B.

APPLICANT: Andrews, William H.

TITLE OF INVENTION: Human Telomerase Catalytic Subunit

NUMBER OF SEQUENCES: 727

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,549A

FILING DATE: 19-NOV-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/911,312

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/912,951

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/915,503

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17618

FILING DATE: 01-OCT-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17885

FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph Ted

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-0026100S

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 214:

SEQUENCE CHARACTERISTICS:

LENGTH: 38 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-974-549A-214

Query Match 62.5%; Score 60; DB 4; Length 38;

Best Local Similarity 83.3%; Pred. No. 0.0019;

Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICQQEYFYFGD 18

Db 1 ICQXEYFYFGD 12

RESULT 8

US-08-851-843A-24

Sequence 24, Application US/08851843A

Patent No. 6093809

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin

APPLICANT: Andrews, William H.

TITLE OF INVENTION: No. 6093809el Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/851,843A

FILING DATE: 06-MAY-1997

CLASSIFICATION:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

CLASSIFICATION:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

CLASSIFICATION:

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-0029300S
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 37 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; US-08-851-843A-24

Query Match 52.6%; Score 50.5; DB 3; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.07;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICQIYYFGDF 18
||| | |||||
Db 1 ICHQ-EYFGDF 11

RESULT 9
US-08-854-050-24
; Sequence 24, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: NO. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643

;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-0029300S
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 37 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; US-08-854-050-24

Query Match 52.6%; Score 50.5; DB 4; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.07;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICQIYYFGDF 18
||| | |||||
Db 1 ICHQ-EYFGDF 11

RESULT 10
US-09-430-323-24
; Sequence 24, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: NO. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:

LENGTH: 39 amino acids

APPLICATION NUMBER: WO PCT/US97/17885

FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002610US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 216:

SEQUENCE CHARACTERISTICS:

LENGTH: 39 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-974-549A-216

Query Match 51.0%; Score 49; DB 4; Length 39;

Best Local Similarity 72.7%; Pred. No. 0.13;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICQIEYYFGD 17

I :I:|||||

Db 1 ILRQVEYYFGD 11

RESULT 13

US-08-854-050-26

Sequence 26, Application US/08854050

Patent No. 6261836

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin

APPLICANT: Andrews, William H.

TITLE OF INVENTION: No. 6261836el Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/854,050

FILING DATE: 09-MAY-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 39 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: peptide

US-08-854-050-26

Query Match 51.0%; Score 49; DB 4; Length 39;

Best Local Similarity 72.7%; Pred. No. 0.13;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 7 ICQIEYYFGD 17

I :I:|||||

Db 1 ILRQVEYYFGD 11

RESULT 14

US-09-430-323-26

Sequence 26, Application US/09430323

Patent No. 6309867

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin

APPLICANT: Andrews, William H.

TITLE OF INVENTION: No. 6309867el Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/430,323

FILING DATE: 29-Oct-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

Search completed: April 23, 2003, 13:36:30
Job time : 12.2247 secs

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-430-323-26

Query Match 51.0%; Score 49; DB 4; Length 39;
Best Local Similarity 72.7%; Pred. NO. 0.13;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICQOIEYFGD 17
| :|:|:|:|:|
Db 1 ILRQVEYFGD 11

RESULT 15
US-08-451-777A-33
; Sequence 33, Application US/08451777A
; Patent No. 5789223
; GENERAL INFORMATION:
; APPLICANT: Bergsma, Derk J.
; APPLICANT: Stambolian, Dwight
; TITLE OF INVENTION: Human Galactokinase Gene
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corp./Corporate
; ADDRESSEE: Intellectual Property
; STREET: 709 Swedeland Road/UW2220
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,777A
; FILING DATE: 26-MAY-1995
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10825
; FILING DATE: 23-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Eagle, Alissa M.
; REGISTRATION NUMBER: 37,126
; REFERENCE/DOCKET NUMBER: P50268-1B
; TELEPHONE: 610-270-5364
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 392 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-451-777A-33

Query Match 45.8%; Score 44; DB 1; Length 392;
Best Local Similarity 57.1%; Pred. NO. 12;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ALEAQICQIEYVF 15
| :|:|:|:|:|
Db 164 ARAQVCCQREHSF 177

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:23 ; Search time 28.5169 Seconds
(without alignments)
84.109 Million cell updates/sec

Title: US-09-836-073-2

Perfect score: 96
Sequence: 1 AALEAQCQIEYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_101002:*
- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
 - 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
 - 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
 - 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
 - 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
 - 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
 - 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
 - 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
 - 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
 - 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
 - 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
 - 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
 - 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
 - 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
 - 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
 - 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
 - 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
 - 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
 - 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
 - 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
 - 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
 - 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
 - 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	90.6	18	21 AAY52200	Human la autoantigen
2	87	90.6	92	21 AAG01351	Human secreted pro
3	87	90.6	408	17 AAW03716	Human autoantigen
4	87	90.6	439	22 ABG08417	Novel human diago
5	87	90.6	460	21 AAB58987	Breast and ovarian
6	87	90.6	460	23 ABP41511	Human ovarian anti
7	65	67.7	21	14 AAR43394	La/SSB epitope 17
8	52	54.2	390	22 ABB65316	Drosophila melanog
9	45	46.9	688	22 ABB61544	Drosophila melanog
10	44	45.8	392	17 AAR87594	Human galactokinas

11	44	45.8	392	19 AAW76730	Human galactokinas
12	44	45.8	392	19 AAW76731	Human galactokinas
13	44	45.8	406	21 AAB58960	Breast and ovarian
14	43	44.8	913	21 AAG47714	Arabidopsis thalia
15	43	44.8	923	21 AAG47713	Arabidopsis thalia
16	43	44.8	993	21 AAG47712	Arabidopsis thalia
17	42	43.8	161	22 AAU32841	Novel human secret
18	42	43.8	211	21 AAB58448	Novel cancer associ
19	42	43.8	328	22 AAB92757	Human protein sequ
20	42	43.8	487	23 ABB98209	Chlamydia polypept
21	42	43.8	582	21 AAY44722	Human immune syste
22	42	43.8	1379	22 ABB68940	Drosophila melanog
23	41	42.7	29	17 AAR87597	Human galactokinas
24	41	42.7	29	19 AAW76729	Human galactokinas
25	41	42.7	229	23 AAU83333	Novel secreted pro
26	41	42.7	382	23 AAU83092	Human polypeptide,
27	41	42.7	493	22 AAG93673	Drosophila melanog
28	40.5	42.2	453	22 ABB60501	Human polypeptide
29	40	41.7	97	22 AAU12729	Human polypeptide
30	40	41.7	262	23 AAU91155	Deinococcus radiop
31	40	41.7	312	22 AAM00871	Human bone marrow
32	40	41.7	378	22 AAM00758	Human bone marrow
33	40	41.7	433	21 AAG34745	Arabidopsis thalia
34	40	41.7	486	23 ABB66706	Human novel polype
35	40	41.7	501	20 AAY05665	Maize cinnamate-4-
36	40	41.7	502	22 ABB25030	Novel human diago
37	40	41.7	502	22 ABB25687	Novel human diago
38	40	41.7	505	15 AAR48205	H. tuberosus cytoch
39	40	41.7	505	15 AAR48206	A. thaliana cytochr
40	40	41.7	505	20 AAY05664	Maize cinnamate-4-
41	40	41.7	505	21 AAG34744	Arabidopsis thalia
42	40	41.7	505	23 AAE20655	Helianthus tuberos
43	40	41.7	505	23 AAE16391	Helianthus tuberos
44	40	41.7	533	21 AAG34743	Arabidopsis thalia
45	40	41.7	647	22 AAB49655	Human SEC7 protein

ALIGNMENTS

RESULT 1

AA132200

ID AAY52200 standard; peptide; 18 AA.

XX AC AAY52200;

XX AC AAY52200;

XX DT 14-MAR-2000 (first entry)

XX DE Human la autoantigen peptide (LAP).

XX DE Human la autoantigen peptide (LAP).

XX KW La autoantigen; LAP; Internal ribosome entry site; IRES; translation; viral replication; IRNA; antiviral agent; picornavirus; flavivirus; coronavirus; hepatitis virus; rhinovirus; adenovirus; coxsackie virus; parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus; foot-and-mouth disease virus; echo virus; infectious bronchitis virus; vesicular stomatitis virus.

XX OS Homo sapiens.

XX PN WO9961613-A2.

XX PD 02-DEC-1999.

XX PF 21-MAY-1999; 99WO-US11281.

XX PR 22-MAY-1998; 98US-0086527.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Das S, Dasgupta A;

XX DR WPI; 2000-062712/05.

XX XX

PT New yeast inhibitory peptide useful for inhibiting viral protein
 translation and replication -
 PS Claim 5; Page 57; 81pp; English.
 XX This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (IRNA see AA45200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rabdovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.

XX Sequence 18 AA;
 Query Match 90.6%; Score 87; DB 21; Length 18;
 Best Local Similarity 88.9%; Pred. No. 1.6e-07;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AALEAQICQIIEYFGDF 18
 |||||:|:| |||||
 Db 1 AALEAKICHQIEYFGDF 18

RESULT 2

AAAG01351
 ID AAG01351 standard; Protein: 92 AA.

XX
 AC AAG01351;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 5432.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.

XX Homo sapiens.

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX N-PSDB; AAC01357.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 13; SEQ ID 5432; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30

CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5',
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 92 AA;

Query Match 90.6%; Score 87; DB 21; Length 92;
 Best Local Similarity 88.9%; Pred. No. 9.6e-07;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AALEAQICQIIEYFGDF 18
 |||||:|:| |||||
 Db 11 AALEAKICHQIEYFGDF 28

RESULT 3

AAW03716

ID AAW03716 standard; protein: 408 AA.

XX
 AC AAW03716;

DT 12-MAR-1997 (first entry)

DE Human autoantigen La(SS-B).

XX Autoimmune disease; La autoantigen; Sjogren's syndrome;
 KW systemic lupus erythematosus; diagnosis.

XX Homo sapiens.

XX US5541291-A.

XX 30-JUL-1996.

XX 31-DEC-1984; 84US-0687908.

XX 27-MAY-1987; 87US-0054871.

XX 31-DEC-1984; 84US-0687908.

XX (UYDU-) UNIV DUKE.

XX Keene JD;

XX WPI; 1996-362015/36.

XX Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
 PT overlap syndrome - useful for diagnosis and treatment of autoimmune
 PT diseases

XX Disclosure; Columns 15-16; 21pp; English.

XX The human lupus antigen (La) is diagnostic for Sjogren's syndrome,
 CC as well as occurring in systemic lupus erythematosus patients. The
 CC La protein is clinically related to the Ro protein that is highly
 CC common among autoimmune patients. La and Ro antigens sometimes
 CC reside on the same cellular ribonucleoprotein particle; most La
 CC patients contain some Ro antibodies and vice versa. La cDNA has
 CC been isolated from a human liver library.

XX Sequence 408 AA;

Query Match 90.6%; Score 87; DB 17; Length 408;
 Best Local Similarity 88.9%; Pred. No. 5e-06;
 Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AALEAQICQIIEYFGDF 18

Db 11 AALEAKICHOIEYFGDF 28
|||||:|||||

RESULT 4
ABG08417
ID ABG08417 standard; Protein: 439 AA.

AC ABG08417;
DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #8408.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN W0200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Dmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS72604.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID No 38776; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantifying a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 439 AA;

Query Match 90.6%; Score 87; DB 22; Length 439;
Best Local Similarity 88.9%; Pred. No. 5.5e-06;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AALEAQICQIEYFGDF 18
|||||:|||||

Db 41 AALEAKICHOIEYFGDF 58

RESULT 5
AAB58987
ID AAB58987 standard; Protein: 460 AA.

XX AAB58987;

XX 27-MAR-2001 (first entry)

DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 695.

XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neoplastic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.

XX Homo sapiens.

XX W0200055173-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05881.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-611515/58.

DR N-PSDB; AAF21890.

XX New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases

PS Claim 11; Page 1149-11150; 1299pp; English.

XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC neoplastic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiac activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemia; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.

XX Sequence 460 AA;

Query Match 90.6%; Score 87; DB 21; Length 460;
Best Local Similarity 88.9%; Pred. No. 5.8e-06;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 AALEAQICQIEYFGDF 18
|||||:|||||

DB 63 AALEAKICHIEYFGDF 80

RESULT 6

ID ABP41511 standard; Protein; 460 AA.

XX AC ABP41511;

XX DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HVVAF56, SEQ ID NO:2643.

XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

KW ovarian cancer; breast cancer; tumour; reproductive system disorder;

KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;

KW inflammatory condition; immune disorder; blood disorder;

KW cardiovascular disorder; respiratory disorder; neurological disorder;

KW gastrointestinal disorder; urinary system disorder; drug screening;

KW gene therapy; chromosome mapping; forensic analysis;

KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;

KW antiinflammatory; gynaecological; reproductive.

XX OS Homo sapiens.

XX PN WO200200677-A1.

XX PD 03-JAN-2002.

XX PF 07-JUN-2001; 2001WO-US18569.

XX PR 07-JUN-2000; 2000US-209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

XX DR N-PSDB; ABQ54588.

XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides, useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian cancer), immune disorders, cardiovascular disorders and neurological diseases -

XX PS Claim 11; SEQ ID No 2643; 2922pp; English.

XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-ABP4328) and to cDNAs encoding them (ABQ54131-ABQ56305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various ovary and/or breast-related disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present

CC sequence represents a human ovarian antigen of the invention.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 460 AA;

Query Match 90.6%; Score 87; DB 23; Length 460;

Best Local Similarity 88.9%; Pred. No. 5.8e-06;

Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AALEAICQCHIEYFGDF 18

Db 63 AALEAKICHIEYFGDF 80

RESULT 7

AAAR43394

ID AAAR43394 standard; peptide; 21 AA.

XX AC AAAR43394;

XX DT 12-MAY-1994 (first entry)

XX DE La/SSB epitope 17.

XX KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD; nuclear ribonucleoprotein; rRNP; Sm B/B'; polypeptide; antigen; D; systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B'; RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

XX OS Homo sapiens.

XX PN WO9321223-A.

XX PD 28-OCT-1993.

XX PF 13-APR-1993; 93WO-US03484.

XX PR 13-APR-1992; 92US-0867819.

XX PA (OKLA) UNIV OKLAHOMA STATE.

XX PI Harley JB;

XX WPI; 1993-351658/44.

XX PT New linear epitope(s) for human auto-antibodies - from the Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used for diagnosing and treating auto-immune disorders e.g. systemic lupus erythematosus

XX PS Claim 1; Page 30; 43pp; English.

XX CC The sequences given in AAAR43391-562 are linear epitopes which are derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen, the 70 kD nuclear ribonucleoprotein (rRNP) and the Sm B/B' polypeptide. These antigens are common in systemic lupus erythematosus (SLE) and closely related disorders. The Ro/SSA family of proteins has been shown to have several molecular forms which are defined by the molecular weight of the antigen identified. The major form has a molecular weight of 60 kD and two additional forms have molecular weights of 52 and 54 kD. La/SSB is also a member of this group of autoantibodies and binds small RNAs with a polyuridine terminus. La/SSB has been shown to be a 46-50 kD monomeric positive sera. La/SSB has been shown to be a 46-50 kD monomeric phosphoprotein which associates with RNA polymerase III transcripts. Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and U5 RNA. Anti-Sm antibodies may be directed against one or a combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD), E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be used for preventing, treating or screening autoimmune disorders, especially SLE or Sjogrens syndrome (SS). They bind to a human

CC autoantibody and may therefore be used as vaccines.

XX Sequence 21 AA;

Query Match 67.7%; Score 65; DB 14; Length 21;
Best Local Similarity 91.7%; Pred. No. 0.00077;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICQIQEYVFGDF 18

Db 1 ICHQIEYVFGDF 12

RESULT 8

ABB65316
ID ABB65316 standard; Protein; 390 AA.

XX Sequence 21 AA;

AC ABB65316;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 22740.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW Pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

PA (PERE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL09419.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

PS Disclosure; SEQ ID NO 22740; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175), and the encoded proteins (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 390 AA;

Query Match 54.2%; Score 52; DB 22; Length 390;

Best Local Similarity 64.3%; Pred. No. 2.7;

Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 EAQICQIEYVFGD 17

Db 51 ERAIRQVEYVFGD 64

RESULT 9

ABB61544
ID ABB61544 standard; Protein; 688 AA.

XX Sequence 21 AA;

AC ABB61544;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 11424.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

PA (PERE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL05647.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

PS Disclosure; SEQ ID NO 11424; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175), and the encoded proteins (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 688 AA;

Query Match 46.9%; Score 45; DB 22; Length 688;

Best Local Similarity 60.0%; Pred. No. 73;

Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 EAQICQIEYVFGDF 18

Db 393 EIQICGIEYVFPFF 407

RESULT 10

AAR87594
ID AAR87594 standard; Protein; 392 AA.

XX Sequence 392 AA;

AC AAR87594;

DT 03-JUN-1996 (first entry)

DE Human galactokinase.

XX Galactokinase; deficiency; galactosemia; diagnosis; therapy;
KW vector; transgenic animal; cataract.

OS Homo sapiens.

XX ZA9407865-A.

XX 29-NOV-1995.

XX 07-OCT-1994; 942A-0007865.

XX 23-SEP-1994; 94WO-US10825.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Bergsma DJ, Stambolian DE;

XX WPI; 1996-049989/05.

XX N-PSDB; AAT08978.

XX

XX New human galacto:kinase gene - useful for galacto:kinase prodn. in

XX host cells and for the diagnosis and treatment of galacto:kinase

XX deficiency in individuals

XX Disclosure; Page 22-25; 3lpp; English.

XX

XX An amino acid sequence for human galactokinase (AAR9594) was

XX deduced from a cDNA clone (AAT08978) isolated from an osteoclastoma

XX stromal cell library. The galactokinase is useful for the

XX treatment of diseases associated with galactokinase deficiency,

XX an autosomal recessive trait with a heterozygote frequency

XX estimated to be 0.2%. The enzyme can also be used to raise

XX antibodies. Recombinant galactokinase is produced by expression

XX of the cDNA in host (prokaryotic or eukaryotic) cells.

XX

XX Sequence 392 AA;

XX

XX Query Match 45.8%; Score 44; DB 17; Length 392;

XX Best Local Similarity 57.1%; Pred. No. 57;

XX Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

XX

XX 2 ALEAQCQIEYF 15

XX | | | | | | | |

XX Db 164 AARQVCQAEHSF 177

XX

XX RESULT 11

XX AAW76730

XX ID AAW76730 standard; Protein; 392 AA.

XX AC AAW76730;

XX 13-JAN-1999 (first entry)

XX Human galactokinase protein.

XX Galactokinase; human; mutation; detection; diagnosis; treatment;

XX deficiency.

XX

XX Homo sapiens.

XX US5830649-A.

XX 03-NOV-1998.

XX

XX 26-MAY-1995; 95US-0451778.

XX

XX 26-MAY-1995; 95US-0451778.

XX

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Bergsma DJ, Stambolian DE;

XX WPI; 1998-609232/51.

XX N-PSDB; AAV62903.

XX

XX Query Match 45.8%; Score 44; DB 19; Length 392;

XX Best Local Similarity 57.1%; Pred. No. 57;

XX

XX Detection of galactokinase mutations - based on comparison with

XX wild-type gene sequence or altered galactokinase activity

XX Claim 1; Column 43-46; 3lpp; English.

XX This sequence represents a novel human galactokinase. This protein is

XX used in a method to detect galactokinase mutations. This protein and its

XX encoding nucleic acid can be used in methods allowing the detection,

XX diagnosis and treatment of human galactokinase deficiency.

XX

XX Sequence 392 AA;

XX

XX Query Match 45.8%; Score 44; DB 19; Length 392;

XX Best Local Similarity 57.1%; Pred. No. 57;

XX Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

XX

XX 2 ALEAQCQIEYF 15

XX | | | | | | | |

XX Db 164 AARQVCQAEHSF 177

XX

XX RESULT 12

XX AAW76731

XX ID AAW76731 standard; Protein; 392 AA.

XX AC AAW76731;

XX 13-JAN-1999 (first entry)

XX Human galactokinase protein variant #1:

XX Galactokinase; human; mutation; detection; diagnosis; treatment;

XX deficiency.

XX

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 32

XX /label= V32M

XX /note= "Wild type Val residue is replaced by Met"

XX

XX US5830649-A.

XX 03-NOV-1998.

XX 26-MAY-1995; 95US-0451778.

XX

XX 26-MAY-1995; 95US-0451778.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Bergsma DJ, Stambolian DE;

XX WPI; 1998-609232/51.

XX N-PSDB; AAV62904.

XX

XX Detection of galactokinase mutations - based on comparison with

XX wild-type gene sequence or altered galactokinase activity

XX

XX Disclosure; Column 23-26; 3lpp; English.

XX

XX This sequence represents a mutant human galactokinase where a Val

XX residue in the wild-type sequence represented in AAW76730 is replaced by

XX a Met. The encoded protein is used in a method to detect galactokinase

XX mutations. This protein and its encoding nucleic acid can be used in

XX methods allowing the detection, diagnosis and treatment of human

XX galactokinase deficiency.

XX

XX Sequence 392 AA;

XX

XX Query Match 45.8%; Score 44; DB 19; Length 392;

XX Best Local Similarity 57.1%; Pred. No. 57;

XX

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ALEAQCQOIEYF 15
| | | | | | | |
Db 164 AARQVCQQAHSF 177

RESULT 13
AAB58960
ID AAB58960 standard; Protein; 406 AA.
XX AC AAB58960;
XX DT 27-MAR-2001 (first entry)
XX DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 668.
XX KW Human: breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; vulnerary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.
XX OS Homo sapiens.
XX PN WO200055173-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US05881.
XX PR 12-MAR-1999; 99US-0124270.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX DR WPI; 2000-611515/58.
XX DR N-FSDB; AAF21863.
XX PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases -
XX PS Claim 11; Page 1116-1118; 1299pp; English.
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antitumor; vulnerary; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiac activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.
XX SQ Sequence 406 AA;

Best Local Similarity 57.1%; Pred. No. 59;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 ALEAQCQOIEYF 15
| | | | | | | |
Db 178 AARQVCQQAHSF 191

RESULT 14
AAG47714
ID AAG47714 standard; Protein; 913 AA.
XX AC AAG47714;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 60168.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 25-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.

```
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140895.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 44.8%; Score 43; DB 21; Length 913;
Best Local Similarity 58.3%; Pred. No. 2.le+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 QICQOIEYFSD 17
Db 116 KIVNQEYFSD 127
```


RESULT 15

AAG47713

ID AAG47713 standard; Protein; 923 AA.

XX AC AAG47713;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 60167.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

XX PR 09-MAR-1999; 99US-0123548.

XX PR 23-MAR-1999; 99US-0125788.

XX PR 25-MAR-1999; 99US-0126264.

XX PR 29-MAR-1999; 99US-0126785.

XX PR 01-APR-1999; 99US-0127462.

XX PR 06-APR-1999; 99US-0128234.

XX PR 08-APR-1999; 99US-0128714.

XX PR 16-APR-1999; 99US-0129845.

XX PR 19-APR-1999; 99US-0130077.

XX PR 21-APR-1999; 99US-0130449.

XX PR 23-APR-1999; 99US-0130510.

XX PR 23-APR-1999; 99US-0130891.

XX PR 28-APR-1999; 99US-0131449.

XX PR 30-APR-1999; 99US-0132048.

XX PR 30-APR-1999; 99US-0132407.

XX PR 04-MAY-1999; 99US-0132484.

XX PR 05-MAY-1999; 99US-0132485.

XX PR 06-MAY-1999; 99US-0132486.

XX PR 07-MAY-1999; 99US-0132863.

XX PR 11-MAY-1999; 99US-0134256.

XX PR 14-MAY-1999; 99US-0134218.

XX PR 14-MAY-1999; 99US-0134219.

XX PR 14-MAY-1999; 99US-0134221.

XX PR 14-MAY-1999; 99US-0134370.

XX PR 18-MAY-1999; 99US-0134768.

XX PR 19-MAY-1999; 99US-0134941.

XX PR 20-MAY-1999; 99US-0135124.

XX PR 21-MAY-1999; 99US-0135353.

XX PR 24-MAY-1999; 99US-0135629.

XX PR 25-MAY-1999; 99US-0136031.

XX PR 27-MAY-1999; 99US-0136392.

XX PR 28-MAY-1999; 99US-0136782.

XX PR 01-JUN-1999; 99US-0137222.

XX PR 03-JUN-1999; 99US-0137528.

XX PR 04-JUN-1999; 99US-0137502.

XX PR 07-JUN-1999; 99US-0137724.

XX PR 08-JUN-1999; 99US-0138094.

XX PR 10-JUN-1999; 99US-0138540.

XX PR 10-JUN-1999; 99US-0138847.

XX PR 14-JUN-1999; 99US-0139119.

XX PR 16-JUN-1999; 99US-0139452.

XX PR 17-JUN-1999; 99US-0139453.

XX PR 17-JUN-1999; 99US-0139492.

XX PR 18-JUN-1999; 99US-0139454.

XX PR 18-JUN-1999; 99US-0139455.

XX PR 18-JUN-1999; 99US-0139456.

XX PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 44.8%; Score 43; DB 21; Length 923;
 Best Local Similarity 58.3%; Pred. No. 2.2e+02;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 QICQOIEYFGD 17
 Db 126 KIVNQEYFSD 137

Search completed: April 23, 2003, 13:27:08
 Job time : 30.5169 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:34:48 ; Search time 20.0225 Seconds
(without alignments)
185.234 Million cell updates/sec

Title: US-09-836-073-2
Perfect score: 96
Sequence: 1 AALPAQCQIYEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 9297

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	30	31.2	24	Q05616	staphylococ
2	29	30.2	16	Q9R963	helicobacte
3	28	29.2	11	Q9RC46	homo sapien
4	27	28.1	18	Q16028	homo sapien
5	27	28.1	25	Q9XJZ7	aureoumbra
6	25.5	26.6	24	Q38270	bacterioph
7	25	26.0	23	Q9UC18	homo sapien
8	25	26.0	24	Q9BM09	spongilla 1
9	25	26.0	25	Q9TGB8	alnus crisp
10	25	26.0	25	Q9TGB7	alnus gluti
11	25	26.0	25	Q9TGB6	alnus marit
12	25	26.0	25	Q9TGB5	betula alle
13	25	26.0	25	Q9TGB4	betula glan
14	25	26.0	25	Q9TGB3	betula papy
15	25	26.0	25	Q9TGB2	betula verr
16	25	26.0	25	Q9TGB1	betula pube

Q9TGB0 corylus ave
Q9TGA9 corylus col
Q9TGA8 corylus cor
Q9TGA7 ostrya virg
Q9TGA6 quercus rub
Q9TGB3 carpinus ca
Q9Z1B4 mus musculu
Q9THR8 bryopsis sp
Q9ELV4 hepatitis b
Q9ELV3 hepatitis b
Q9ELV2 hepatitis b
Q9ELV1 hepatitis b
Q9ELV0 hepatitis b
Q9ELU9 hepatitis b
Q9ELU8 hepatitis b
Q9ELU6 hepatitis b
Q9ELU5 hepatitis b
Q9ELU4 hepatitis b
Q9ELU3 hepatitis b
Q9ELU2 hepatitis b
Q9ELU1 hepatitis b
Q9ELU0 hepatitis b
Q9ELT9 hepatitis b
Q9ELT8 hepatitis b
Q45663 bacillus su
Q9UCE8 homo sapien
Q96BE3 homo sapien
Q9TWR0 blattella g
Q9S8X5 glycine max

ALIGNMENTS

RESULT 1

Q05616 PRELIMINARY; PRT; 24 AA.
AC Q05616;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-NOV-1998 (TREMREL. 08, Last annotation update)
DE 3-dehydroquinase synthase (EC 4.6.1.3) (Fragment).
GN AROB.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8325-4;
RX MEDLINE=93381456; PubMed=8371108;
RA O'Connell C.M., Patter P., Foster T.J.;
RT "Sequence and mapping of the araA gene of Staphylococcus aureus 8325-4."
RT J. Gen. Microbiol. 139:1449-1460(1993).
RL J.
CC -1- CATALYTIC ACTIVITY: 7-PHOSPHO-3-DEOXY-ARABINO-HEPTULOSONATE - 3-DEHYDROQUINATE + ORTHOPHOSPHATE.
CC -1- PATHWAY: SECOND STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
DR EMBL; L05004; AAA71896.1;
KW Aromatic amino acid biosynthesis; Lyase.
FT NON_TER 1
SQ SEQUENCE 24 AA; 2857 MW; 1C4C8D20F7A4856E CRC64;

Query Match 31.2%; Score 30; DB 2; Length 24;
Best Local Similarity 50.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 CQOIYYF 15

DB 16 CEQLKTYF 23

```

RESULT 2
Q9R963 PRELIMINARY; PRT; 16 AA.
ID Q9R963;
AC Q9R963;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE RepC (Fragment).
GN FEPC.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F31;
RX MEDLINE=98453456; PubMed=9780260;
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,
RA Kuriyama M.;
RT "Full-length sequence analysis of the vacA gene from cytotoxic and
RT nontoxic Helicobacter pylori."
RL J. Infect. Dis. 178:1391-1398(1998).
DR EMBL; AF049623; AAD04263.1; -
FT NON_TER 1
FT NON_TER 16
FT NON_TER 16
SQ SEQUENCE 16 AA; 1938 MW; 3C33D403AEE61428 CRC64;

Query Match 30.2%; Score 29; DB 2; Length 16;
Best Local Similarity 55.6%; Pred. No. 3.9e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYFGDF 18
|:|:|:|:|
Db 7 QVEYAFNF 15

RESULT 3
Q9UC46 PRELIMINARY; PRT; 11 AA.
ID Q9UC46;
AC Q9UC46;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Neutrophil inhibitor peptide, NIP-POLYNORPHONUCLEAR neutrophil
DE inhibitor peptide.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=96326114; PubMed=8703476;
RA Cooper J.A.Jr., Culbreth R.R.;
RT "Characterization of a neutrophil inhibitor peptide harvested from
RT human bronchial lavage: homology to influenza A nucleoprotein."
RL Am. J. Respir. Cell Mol. Biol. 15:207-215(1996).
SQ SEQUENCE 11 AA; 1262 MW; 951A1C3279C9DB45 CRC64;

Query Match 29.2%; Score 28; DB 4; Length 11;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 13 YFGD 17
|:|:|:|
Db 5 YFGD 9

RESULT 4
Q16028 PRELIMINARY; PRT; 18 AA.
ID Q16028;
AC Q16028;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

```

```

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE OCRL-1 protein (Fragment).
GN OCRL-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93278398; PubMed=8504307;
RA Leahy A.M., Charnas L.R., Nussbaum R.L.;
RT "Nonsense mutations in the OCRL-1 gene in patients with the
RT oculocerebrorenal syndrome of Lowe."
RL Hum. Mol. Genet. 2:461-463(1993).
DR EMBL; S62084; AAD13933.1; -
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 18 AA; 2055 MW; 61356C825F7AF33D CRC64;

Query Match 28.1%; Score 27; DB 4; Length 18;
Best Local Similarity 36.4%; Pred. No. 9.7e+02;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALEAQICQOI 11
|:|:|:|:|
Db 1 SAYDPRICROL 11

RESULT 5
Q9XJZ7 PRELIMINARY; PRT; 25 AA.
ID Q9XJZ7;
AC Q9XJZ7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit
DE (Fragment).
GN RBCS.
OS Aureobrombura lagunensis.
OG Chloroplast.
OC Eukaryota; stramenopiles; Pelagophyceae; Aureobrombura.
OX NCBI_TaxID=44058;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCMP1681, CCMP1509, AND CCMP1510;
RA Bailey J.C., Andersen R.A.;
RT "Analysis of clonal cultures of the brown tide algae Aureococcus and
RT Aureobrombura (Pelagophyceae) using 18S rRNA, rbcL and rubisco spacer
RT sequences."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF117782; AAD39106.1; -
DR EMBL; AF117780; AAD39104.1; -
DR EMBL; AF117781; AAD39105.1; -
DR InterPro; IPR000894; RuBisCO_small.
DR Pfam; PF00101; RuBisCO_small; 1.
KW Chloroplast.
FT NON_TER 25
FT NON_TER 25
SQ SEQUENCE 25 AA; 2972 MW; 61A96BD83DB94118 CRC64;

Query Match 28.1%; Score 27; DB 8; Length 25;
Best Local Similarity 62.5%; Pred. No. 1.4e+03;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 6 QICQIQIEY 13
|:|:|:|
Db 18 QIRKQVEY 25

RESULT 6
Q38270 PRELIMINARY; PRT; 24 AA.
ID Q38270;
AC Q38270;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

```

DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Bacteriophage lambda mutant truncated regulatory protein (PyrI) gene
 DE from E.coli, 3' end (Fragment).
 OS Bacteriophage lambda.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
 OC Lambda-like viruses.
 OX NCBI_TaxID=10710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86143826; PubMed=3912513;
 RA Cunin R., Jacobs A., Charlier D., Crabeel M., Herve G., Glansdorff N.,
 RA Pierard A.: function relationship in allosteric aspartate
 RT carbamoyltransferase from Escherichia coli: I. Primary structure of a
 RT pyrI gene encoding a modified regulatory subunit.";
 RL J. Mol. Biol. 186:707-713(1985).
 DR EMBL; M28579; AAA32252.1;
 DR HSSP; P00478; 8ATC.
 FT NON_TER 1
 SQ SEQUENCE 24 AA; 2905 MW; FD9349DF6F8159D0 CRC64;

Query Match 26.6%; Score 25.5; DB 9; Length 24;
 Best Local Similarity 38.5%; Pred. No. 2.3e+03;
 Matches 5; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

QY 2 ALEAICQIQEY 14
 II: : : : :
 DB 8 ALKCKYCEK-EFY 19

RESULT 7

Q9UCL8
 ID Q9UCL8 PRELIMINARY; PRT; 23 AA.
 AC Q9UCL8;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
 DE GLYCOSAMINOGLYCAN-BEARING POLYPEPTIDE (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=93098810; PubMed=1463459;
 RA Bonnet F., Perin J.P., Maillet P., Jolles P., Alliel P.M.;
 RT "Characterization of a human seminal plasma glycosaminoglycan-bearing
 FT polypeptide.";
 RL Biochem. J. 288:565-569(1992).
 SQ SEQUENCE 23 AA; 2334 MW; E469397E6B2EA794 CRC64;

Query Match 26.0%; Score 25; DB 4; Length 23;
 Best Local Similarity 36.4%; Pred. No. 2.7e+03;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 CQIQEYFGDF 18
 I:: : : :
 DB 7 CEEQETXGDF 17

RESULT 8

Q9BM09
 ID Q9BM09 PRELIMINARY; PRT; 24 AA.
 AC Q9BM09;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DE Gypsy-like reverse transcriptase (Fragment).
 OS Spongilla lacustris (Freshwater sponge).
 OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;
 OC Haplosclerida; Spongillidae; Spongilla.
 OX NCBI_TaxID=6055;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TRANSPOSON-GRT-G7 RETROTRANSPOSON;
 RX MEDLINE=20570504; PubMed=11121049;
 RA Arkhipova I., Meseison M.;
 RT "Transposable elements in sexual and asexual taxa.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
 DR EMBL; AY013997; AAG59969.1;
 KW RNA-directed DNA polymerase.
 FT NON_TER 1
 FT NON_TER 24
 SQ SEQUENCE 24 AA; 2754 MW; B588512EDE7A251A CRC64;

Query Match 26.0%; Score 25; DB 5; Length 24;
 Best Local Similarity 33.3%; Pred. No. 2.8e+03;
 Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 ICQIQEYF 15
 I: : : : :
 DB 15 VCHGLEFVF 23

RESULT 9

Q9TGB8
 ID Q9TGB8 PRELIMINARY; PRT; 25 AA.
 AC Q9TGB8;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Ribosomal protein small 3 (Fragment).
 GN RPS3.
 OS Alnus crispa.
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fagales; Betulaceae; Alnus.
 OX NCBI_TaxID=3518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99261656; PubMed=10331271;
 RA Laroche J., Bousquet J.;
 RT "Evolution of the mitochondrial rps3 intron in perennial and annual
 RT angiosperms and homology to nad5 intron 1.";
 RL Mol. Biol. Evol. 16:441-452(1999).
 DR EMBL; AF080075; AAD50062.1;
 DR InterPro; IPR001351; Ribosomal_S3.
 DR Pfam; PF00417; Ribosomal_S3_N; 1.
 KW Mitochondrion.
 FT NON_TER 1
 FT NON_TER 25
 SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 26.0%; Score 25; DB 8; Length 25;
 Best Local Similarity 60.0%; Pred. No. 2.9e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFVG 16
 I:: : : :
 DB 4 DYVG 8

RESULT 10

Q9TGB7
 ID Q9TGB7 PRELIMINARY; PRT; 25 AA.
 AC Q9TGB7;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE Ribosomal protein small 3 (Fragment).
 GN RPS3.
 OS Alnus glutinosa (Alder).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=3517;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99261656; PubMed-10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RL Mol. Biol. Evol. 16:441-452(1999).
RL angiosperms and homology to nad5 intron 1.";
DR EMBL; AF080076; AAD50063.1; -
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;
Query Match 26.0%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.9e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 12 EYFYG 16
Db :||:|
4 DYYYG 8
RESULT 11
Q9TGB6 PRELIMINARY; PRT; 25 AA.
AC Q9TGB6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN rps3.
OS Alnus maritima.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=21015;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99261656; PubMed-10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RL Mol. Biol. Evol. 16:441-452(1999).
RL angiosperms and homology to nad5 intron 1.";
DR EMBL; AF080077; AAD50064.1; -
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;
Query Match 26.0%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.9e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 12 EYFYG 16
Db :||:|
4 DYYYG 8
RESULT 12
Q9TGB5 PRELIMINARY; PRT; 25 AA.
AC Q9TGB5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Betula alleghaniensis.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=21017;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99261656; PubMed-10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RL Mol. Biol. Evol. 16:441-452(1999).
RL angiosperms and homology to nad5 intron 1.";
DR EMBL; AF080078; AAD50065.1; -
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;
Query Match 26.0%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.9e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 12 EYFYG 16
Db :||:|
4 DYYYG 8
RESULT 13
Q9TGB4 PRELIMINARY; PRT; 25 AA.
AC Q9TGB4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Betula glandulosa.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=21018;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99261656; PubMed-10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RL Mol. Biol. Evol. 16:441-452(1999).
RL angiosperms and homology to nad5 intron 1.";
DR EMBL; AF080079; AAD50066.1; -
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;
Query Match 26.0%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.9e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 12 EYFYG 16
Db :||:|
4 DYYYG 8
RESULT 14
Q9TGB3

Db 4 DYYG 8

Search completed: April 23, 2003, 13:47:13
Job time : 21.0225 secs

ID Q9TGB3 PRELIMINARY; PRT; 25 AA.
AC Q9TGB3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Betula papyrifera (Paper birch).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=3507;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RT angiosperms and homology to nad5 intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080080; AAD50067.1; -;
DR InterPro: IPR001351; Ribosomal_S3.
DR Pfam: PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;
Query Match 26.0%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.9e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 12 EYVFG 16
Db :||:|
4 DYYG 8

RESULT 15
Q9TGB2
ID Q9TGB2 PRELIMINARY; PRT; 25 AA.
AC Q9TGB2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Betula verrucosa (White birch) (Betula pendula).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=3505;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RT angiosperms and homology to nad5 intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080081; AAD50068.1; -;
DR InterPro: IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;
Query Match 26.0%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.9e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 12 EYVFG 16
Db :||:|

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:33:13 ; Search time 4.65169 Seconds
(without alignments)
160.495 Million cell updates/sec

Title: us-09-836-073-2
Perfect score: 96
Sequence: 1 AALEAQICQIIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 1520

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	28.1	19	1 HBB2_UROHA	P18992 uromastix h
2	24	25.0	16	1 MLB_SQAC	P01207 squalus aca
3	24	25.0	18	1 MLB_SQAC	P01207 scyllorhinu
4	24	25.0	20	1 PNV2_PHONI	Q5twr5 phoneutria
5	23	24.0	25	1 ANDT_ANDAU	P56684 androctonus
6	22	22.9	18	1 OBP_LYMDI	P34173 lymantria d
7	22	22.9	20	1 COXN_THUOB	P80980 thunnus obe
8	22	22.9	25	1 SMBP_RAT	P80968 rattus norv
9	21	21.9	23	1 NUO5_SOLTU	P80262 solanum tub
10	21	21.9	24	1 HS9B_RABIT	P30947 cryctolagus
11	20.5	21.4	23	1 UDP_IACCA	P19662 lactobacill
12	20	20.8	18	1 SFAH_HELAN	P81098 hellanthus
13	20	20.8	19	1 ADC_CLOPA	P81336 clostridium
14	20	20.8	20	1 COG4_CHIOP	P34156 chionoecete
15	20	20.8	20	1 FIBB_FELCA	P14469 felis silve
16	20	20.8	23	1 CLVB_STVCL	P80711 styela clav
17	20	20.8	23	1 PRO3_DAGCL	P18690 dactylis gl
18	20	20.8	23	1 XYCL_ACIGB	P46365 acinetobact
19	19	19.8	8	1 AL17_CARMA	P81820 carcinus ma
20	19	19.8	9	1 LPF_ECOLI	P30577 escherichia
21	19	19.8	15	1 DHE2_THUTH	P20016 thunnus thy
22	19	19.8	16	1 FIBB_VULVU	P14482 vulpes vulp
23	19	19.8	19	1 FLA2_SPIAU	P21985 spirochaeta
24	19	19.8	20	1 THIO_CANFA	P99505 canis fami
25	19	19.8	20	1 TL19_ARATH	P82658 arabidopsis
26	19	19.8	21	1 LCA_FELCA	P37154 felis silve
27	19	19.8	24	1 PRLA_ACHLY	P27459 achromobact
28	19	19.8	25	1 RL36_MYCCA	Q48972 mycoplasma
29	19	19.8	25	1 OCPI_OCTMI	P58648 octopus min
30	18	18.8	4	1 LMT2_LOCFI	P22396 locusta stig
31	18	18.8	8	1 NP1_LYMST	P80178 lymnaea sta
32	18	18.8	13	1 NP2_LYMST	P80179 lymnaea sta
33	18	18.8	13	1 NP3_LYMST	P80180 lymnaea sta
				1 NP4_LYMST	P80181 lymnaea sta
				1 NP5_LYMST	P80182 lymnaea sta
				1 KLP5_SCARA	P58396 scaptocosa
				1 LECB_POSC	P22584 psophocarpu
				1 FIBA_CERSI	P14535 ceratotheri
				1 FIBA_MACFU	P12803 macaca fusc
				1 FIBA_MANDRILLUS	P14455 mandrillus
				1 FIBA_MANLE	P81666 pinus pinas
				1 TPIS_PINPS	P50982 conus ermin
				1 CXAL_CONER	P34172 hyalophora
				1 OBP_HYACE	P14441 bison bonas
				1 FIBA_BISBO	

RESULT 1

ID	HBB2_UROHA	STANDARD;	PRT;	19 AA.
AC	P18992;			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Hemoglobin beta-2 chain (Fragment).			
OS	Uromastix hardwickii (indian spiny-tailed lizard).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Uromastycinae;			
OC	Uromastix.			
OX	NCBI_TaxID=40250;			
RN	[1]			
RP	SEQUENCE.			
RX	MDLINE=84029159; PubMed=6628672;			
RA	Naqvi S., Zaidi Z.H., von Bahr-Lindstroem H., Carlquist M.,			
RA	Joernvall H.;			
RT	"Characterization of hemoglobin from the lizard Uromastix			
RT	hardwickii.";			
RL	FEBS Lett. 162:290-295(1983).			
CC	-1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE			
CC	VARIOUS PERIPHERAL TISSUES.			
CC	-1- SUBUNIT: HETEROTETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.			
CC	-1- TISSUE SPECIFICITY: RED BLOOD CELLS.			
CC	-1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.			
DR	PIR; A05305; A05305.			
DR	InterPro: IPR000971; Globin.			
DR	PROSITE: PS01033; GLOBIN; PARTIAL.			
KW	Heme; Oxygen transport; Transport; Erythrocyte.			
FT	NON_TER 1			
FT	NON_TER 19			
SQ	SEQUENCE 19 AA; 1914 MW; C40AD8EA30019057 CRC64;			

Query Match 28.1%; Score 27; DB 1; Length 19;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY	14 YFGDF 18
DB	1 FFGDF 5

RESULT 2

ID	MLB_SQAC	STANDARD;	PRT;	16 AA.
AC	P01207;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Melanotropin beta.			
OS	Squalus acanthias (Spiny dogfish).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;			
OC	Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.			
OC	NCBI_TaxID=7797;			
RN	[1]			


```

RP SEQUENCE.
RX MEDLINE=75127390; PubMed=4375978;
RA Bennett H.P.J., Lowry P.J., McMartin C., Scott A.P.;
RT "Structural studies of alpha-melanocyte-stimulating hormone and a
RT novel beta-melanocyte-stimulating hormone from the neurointermediate
RL lobe of the pituitary of the dogfish Squalus acanthias.";
RL Blochem. J. 141:439-444(1974).
CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
DR PIR; A01471; MTDFBS.
KW Hormone.
SQ SEQUENCE 16 AA; 1939 MW; 993AF43C8A640A0E CRC64;

Query Match 25.0%; Score 24; DB 1; Length 16;
Best Local Similarity 57.1%; Pred. No. 3.7e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 12 EYFEGDF 18
:| | |
Db 4 DYKFGHF 10

RESULT 3
MLB_SCYCA STANDARD; PRT; 18 AA.
ID MLB_SCYCA AC P01206;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanotropin beta (Beta-MSH).
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyllorhinidae; Scyllorhinus.
OX NCBI_TaxID=7830;
RN [1]
RN SEQUENCE.
RX MEDLINE=75113445; PubMed=4452470;
RA Love R.M., Pickering B.T.;
RT "A beta-MSH in the pituitary gland of the spotted dogfish
RT (Scyllorhinus canicula): isolation and structure.";
RL Gen. Comp. Endocrinol. 24:398-404(1974).
CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
DR PIR; A01470; MTDFBC.
KW Hormone.
SQ SEQUENCE 18 AA; 2195 MW; 353C93A30C314DF4 CRC64;

Query Match 25.0%; Score 24; DB 1; Length 18;
Best Local Similarity 44.4%; Pred. No. 4.3e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 QIEYFEGDF 18
:| | |
Db 2 ZIBYKMGHF 10

RESULT 4
PNV2_PHONI STANDARD; PRT; 20 AA.
ID PNV2_PHONI AC Q9TWR5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Toxin PNV2 (Fragment).
OS Phoneutria nigriventer (Brazilian armed spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.
OX NCBI_TaxID=6918;
RN [1]
RN SEQUENCE.
RX TISSUE=Venom;
SQ MEDLINE=94030062; PubMed=8216354;
RA Bento A.C., Novello J.C., Marangoni S., Antunes E., Giglio J.R.,
RA Oliveira B., de Nucci G.;

```

```

RT "Identification of a new vascular smooth muscle contracting
RT polypeptide in Phoneutria nigriventer spider venom.";
RL Biochem. Pharmacol. 46:1092-1095(1993).
CC -1- FUNCTION: Has a vascular smooth muscle contracting activity.
CC Causes short-lived contractions of both arterial and venous rabbit
CC vessels.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -1- PTM: MAY POSSESS FOUR DISULFIDE BONDS.
KW Toxin.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2176 MW; F28C3D81D983BCA5 CRC64;

Query Match 25.0%; Score 24; DB 1; Length 20;
Best Local Similarity 62.5%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALPAQICQ 9
:| | |
Db 2 AKRADICQ 9

RESULT 5
ANDT_ANDAU STANDARD; PRT; 25 AA.
ID ANDT_ANDAU AC P56684; P81616;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Androctonin.
OS Androctonus australis hector (Sahara scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Butiidae; Butiidae; Androctonus.
OX NCBI_TaxID=70175;
RN [1]
RN SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
RX TISSUE=Hemolymph;
RX MEDLINE=97094646; PubMed=8939880;
RA Ehret-Sabatier L., Loew D., Goyffon M., Fehlbaum P., Hoffmann J.A.,
RA van Dorsael A., Bulet P.;
RT "Characterization of novel cysteine-rich antimicrobial peptides from
RT scorpion blood.";
RL J. Biol. Chem. 271:29537-29544(1996).
RN [2]
RN SYNTHESIS OF D-AMINO ACID ENANTIOMER, AND CHARACTERIZATION.
RX MEDLINE=20115101; PubMed=10642525;
RA Hetru C., Letellier L., Oren Z., Hoffmann J.A., Shai Y.;
RT "Androctonin, a hydrophilic disulphide-bridged non-haemolytic
RT anti-microbial peptide: a plausible mode of action.";
RL Biochem. J. 345:653-664(2000).
RN [3]
RN STRUCTURE BY NMR.
RX MEDLINE=20025109; PubMed=10563585;
RA Mandard N., Sy D., Maufrais C., Bonmatin J.M., Bulet P., Hetru C.,
RA Vovelle F.;
RT "Androctonin, a novel antimicrobial peptide from scorpion Androctonus
RT australis: solution structure and molecular dynamics simulations in
RT the presence of a lipid monolayer.";
RL J. Biomol. Struct. Dyn. 17:367-380(1999).
CC -1- FUNCTION: ACTIVE AGAINST BOTH BACTERIA (GRAM-POSITIVE AND GRAM-
CC NEGATIVE) AND FILAMENTOUS FUNGI. ACTS ON THE MEMBRANE OF THE
CC BACTERIAL CELLS. IT DESTABILIZE A MEMBRANE BY MODIFYING ITS
CC PROPERTIES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MASS SPECTROMETRY: MW=3076.7; METHOD=Electrospray.
DR PDB; 1CZ6; 12-JAN-00.
KW Antibiotic; Fungicide; 3D-structure.
FT DISULFID 4 20
FT DISULFID 10 16
SQ SEQUENCE 25 AA; 3081 MW; D4183D6EA5AA50AA CRC64;

Query Match 24.0%; Score 23; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 9.1e+02;

```

Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICQIE 12
: : : :
Db 3 VCRQIK 8

RESULT 6
OBP_LYMDI STANDARD; PRT; 18 AA;
AC P34173;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE General odorant-binding protein (GOBP) (Fragment).
OS Lymantria dispar (Gypsy moth).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC Ditrysia; Noctuoidea; Lymantrilidae; Lymantria.
OX NCBI_TaxID=13123;
RN [1]
RP SEQUENCE.
RX MEDLINE-91186129; PubMed-2010751;
RA Vogt R.G., Prestwich G.D., Lerner M.R.;
RT "Odorant-binding-protein subfamilies associate with distinct classes
of olfactory receptor neurons in insects.";
RL J. Neurobiol. 22:74-84(1991).
CC -1- FUNCTION: PRESENT IN THE AQUEOUS FLUID SURROUNDING OLFACTORY
SENSORY DENDRITES AND ARE THOUGHT TO AID IN THE CAPTURE AND
TRANSPORT OF HYDROPHOBIC ODORANTS INTO AND THROUGH THIS FLUID.
CC -1- SUBUNIT: HOMODIMER (PROBABLE).
CC -1- TISSUE SPECIFICITY: ANTENNA.
CC -1- SIMILARITY: BELONGS TO THE PBP/GOBP FAMILY.
KW Olfaction; Transport.
FT NON_TER 18 18
SQ SEQUENCE 18 AA; 1966 MW; 88A1B38FDB8021A CRC64;

Query Match 22.9%; Score 22; DB 1; Length 18;
Best Local Similarity 25.08; Pred. No. 9.5e+02;
Matches 3; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 5 AOICQIEYFG 16
: : : :
Db 2 AEVMSVXAHFG 13

RESULT 7
COXN_THUOB STANDARD; PRT; 20 AA.
AC P80980;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Cytochrome c oxidase polypeptide Viib-heart (EC 1.9.3.1) (Fragment).
OS Thunus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE.
RX MEDLINE-97454291; PubMed-9310366;
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
liver.";
RL Eur. J. Biochem. 248:99-103(1997).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
MITOCHONDRIAL ELECTRON TRANSPORT.
CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) -> 4 ferricytochrome

CC c + 2 H(2)O.
KW Oxidoreductase; Mitochondrion.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2303 MW; 0A33BD34006E5AA6 CRC64;

Query Match 22.9%; Score 22; DB 1; Length 20;
Best Local Similarity 33.3%; Pred. No. 1.1e+03;
Matches 3; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 QQIEYFGD 17
: : : :
Db 6 QDPHLYGD 14

RESULT 8
SMBP_RAT STANDARD; PRT; 25 AA.
ID SMBP_RAT
AC P80968;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SM-11044 binding protein (Fragments).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RX STRAIN-Wistar;
RX MEDLINE-97407910; PubMed-9261134;
RA Sugawara T., Matsuzaki-Fujita M., Guillaume J.-L., Camoin L.,
RA Morooka S., Strosberg A.D.;
RT "Characterization of a novel iodoctanopindolol and SM-11044 binding
protein, which may mediate relaxation of depolarized rat colon
tonus.";
RT J. Biol. Chem. 272:21244-21252(1997).
CC -1- FUNCTION: MAY MEDIATE RELAXATION OF DEPOLARIZED COLON TONUS.
IT BINDS IODOCTANOPINDOLOL AND SM-11044.
CC -1- MISCELLANEOUS: THE ORDER OF THE FRAGMENTS IS UNKNOWN.
FT NON_TER 1 1 OR Y.
FT UNSURE 6 6
FT NON_CONS 18 19
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3177 MW; D14F0CB9B778C2CB CRC64;

Query Match 22.9%; Score 22; DB 1; Length 25;
Best Local Similarity 28.6%; Pred. No. 1.4e+03;
Matches 4; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 5 AOICQIEYFGDF 18
: : : :
Db 8 ADPARYFQYFPXF 21

RESULT 9
NUO5_SOLTU STANDARD; PRT; 23 AA.
ID NUO5_SOLTU
AC P80262;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase 29 kDa subunit (EC 1.6.5.3)
DE (EC 1.6.99.3) (Complex I-29KD) (CI-29KD) (Fragment).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RX STRAIN-cv. Bintje; TISSUE-Tuber;
RX MEDLINE-94124587; PubMed-8294484;
RA Herz U., Schroeder W., Liddell A., Leaver C.J., Brennicke A.,

RA Grohmann L.;
 RT "Purification of the NADH:ubiquinone oxidoreductase (complex I) of
 the respiratory chain from the inner mitochondrial membrane of
 Solanum tuberosum.";
 RL J. Biol. Chem. 269:2263-2269(1994).
 CC -|- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
 CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
 TO BE UBIQUINONE.
 CC -|- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -|- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
 CC -|- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
 CC -|- SUBCELLULAR LOCATION: MATRIX SIDE OF THE MITOCHONDRIAL INNER
 MEMBRANE.
 DR PIR: C49732; C49732.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2653 MW; 4B41B8F83412F58 CRC64;
 Query Match 21.9%; Score 21; DB 1; Length 23;
 Best Local Similarity 40.0%; Pred. No. 1.8e+03;
 Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 12 EYFYG 16
 DB 18 DYFYG 22
 RESULT 10
 HS9B_RABIT
 ID HS9B_RABIT STANDARD; PRT; 24 AA.
 AC P30947;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Heat shock protein HSP 90-beta (HSP 84) (Fragment).
 GN HSPCB.
 OS Oryctolagus cuniculus (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=90008887; PubMed=2507541;
 RA Lees-Miller S., Anderson C.W.;
 RT "The human double-stranded DNA-activated protein kinase phosphorylates
 the 90-kDa heat-shock protein, hsp90 alpha at two NH2-terminal
 threonine residues.";
 RL J. Biol. Chem. 264:17275-17280(1989).
 CC -|- FUNCTION: MOLECULAR CHAPERONE. HAS ATPASE ACTIVITY
 (BY SIMILARITY).
 CC -|- SUBUNIT: Homodimer (By similarity).
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -|- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
 DR PIR: B34461; B34461.
 DR HSP; P07900; IBIQ.
 DR InterPro; IPR001404; Hsp90.
 DR PROSITE; PS00298; HSP90; PARTIAL.
 KW Chaperone; ATP-binding; Heat shock; Phosphorylation.
 FT NON_TER 24 24
 SQ SEQUENCE 24 AA; 2771 MW; 0FE0716C6948FC33 CRC64;
 Query Match 21.9%; Score 21; DB 1; Length 24;
 Best Local Similarity 50.0%; Pred. No. 1.9e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 ALEAQICQ 9
 DB 15 AFQAEIAQ 22
 RESULT 11
 UDP_LACCA
 ID UDP_LACCA STANDARD; PRT; 23 AA.

P19662;
 AC 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Uridine phosphorylase (EC 2.4.2.3) (UDRPase) (Fragment).
 GN UDP.
 OS Lactobacillus casei.
 CC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 CC Lactobacillus.
 OX NCBI_TaxID=1582;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=ATCC 7469;
 RX MEDLINE=90381286; PubMed=2119230;
 RA Avraham Y., Grossowicz N., Yashphe J.;
 RT "Purification and characterization of uridine and thymidine
 phosphorylase from Lactobacillus casei.";
 RL Biochim. Biophys. Acta 1040:287-293(1990).
 CC -|- FUNCTION: THE ENZYMES WHICH ARE INVOLVED IN THE DEGRADATION OF THESE
 OF PYRIMIDINE NUCLEOSIDES ARE INVOLVED IN THE UTILIZATION AS CARBON AND ENERGY SOURCES,
 CC COMPOUNDS AND IN THEIR UTILIZATION AS CARBON AND ENERGY SOURCES.
 CC OR IN THE RESCUE OF PYRIMIDINE BASES FOR NUCLEOTIDE SYNTHESIS.
 CC -|- CATALYTIC ACTIVITY: Uridine + phosphate = uracil + alpha-D-ribose
 1-phosphate.
 CC -|- PATHWAY: Nucleotide and deoxyribonucleotide catabolism.
 CC -|- SUBUNIT: HOMOTETRAMER.
 CC -|- SIMILARITY: BELONGS TO THE PNP/UDP FAMILY 1 OF PHOSPHORYLASES.
 DR PIR: S11383; S11383.
 DR InterPro; IPR000845; PNP_UDP.
 DR PROSITE; PS01232; PNP_UDP_1; PARTIAL.
 KW Transferrase; Glycosyltransferase.
 FT UNSURE 1 1 OR S.
 FT UNSURE 7 7 OR T.
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2630 MW; 918B2E2F32F35A17 CRC64;
 Query Match 21.4%; Score 20.5; DB 1; Length 23;
 Best Local Similarity 36.4%; Pred. No. 2.2e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 3; Gaps 1;
 QY 7 ICQIIEYFGD 17
 DB 15 VCTDV---FGD 22
 RESULT 12
 SFAH_HELAN
 ID SFAH_HELAN STANDARD; PRT; 18 AA.
 AC P81098;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Seed fatty acyl-ester hydrolase (EC 3.1.1.1) (Fragment).
 OS Helianthus annuus (Common sunflower).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; euasterids II; Asterales; Asteraceae; Asteroidae;
 CC Helianthaceae; Helianthus.
 OX NCBI_TaxID=4232;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CV. RUSTICA / VAR. EUROFLOR; TISSUE=Seed;
 RA Beisson F., Gardies A.-M., Teissere M., Ferte N., Noat G.;
 RT "An esterase neosynthesized in post-germinated sunflower seeds is
 related to a new family of lipolytic enzymes.";
 RL Plant Physiol. Biochem. 35:761-765(1997).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=95210327; PubMed=7696323;
 RA Teissere M., Borel M., Caillol B., Nari J., Gardies A.-M., Noat G.;
 RT "Purification and characterization of a fatty acyl-ester hydrolase
 from post-germinated sunflower seeds.";
 RL Biochim. Biophys. Acta 1255:105-112(1995).

```
CC -1- FUNCTION: IMPLICATED IN THE BREAKDOWN OF OIL BODY-STORED LIPIDS
CC DURING POST-GERMINATION.
CC -1- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O = an alcohol + a
CC carboxylic anion.
CC -1- TISSUE SPECIFICITY: SEED.
CC -1- DEVELOPMENTAL STAGE: POST-GERMINATION.
CC -1- SIMILARITY: BELONGS TO THE "GDSL" FAMILY OF LIPOLYTIC ENZYMES.
DR InterPro: IPR001087; Lipase_GDSL.
DR PROSITE: PS01098; LIPASE_GDSL_SER; PARTIAL.
KW Hydrolase; Lipid degradation; Glycoprotein.
FT ACT_SITE 13 BY SIMILARITY.
FT NON_TER 18
SQ SEQUENCE 18 AA; 2057 MW; 3E9A80EB548E862 CRC64;

Query Match 20.8%; Score 20; DB 1; Length 18;
Best Local Similarity 60.0%; Pred. No. 2.1e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 YFGD 17
DB 8 FFGD 12

RESULT 13
ADC_CLOPA STANDARD; PRT; 19 AA.
AC P81336;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Acetoacetate decarboxylase (EC 4.1.1.4) (ADC) (CP 28/CP 29)
DE (Fragment).
GN ADC.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]
RP SEQUENCE.
RC STRAIN=W5.
RX MEDLINE=98291870; PubMed=9629918;
RA Flengsrud R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5.";
RL Electrophoresis 19:802-806(1998).
CC -1- CATALYTIC ACTIVITY: Acetoacetate + H(+) = acetone + CO(2).
CC -1- SUBUNIT: HOMODODECAMER (BY SIMILARITY).
KW Lyase; Decarboxylase.
FT NON_TER 19
SQ SEQUENCE 19 AA; 2078 MW; 6ACDAA91103AEC31 CRC64;

Query Match 20.8%; Score 20; DB 1; Length 19;
Best Local Similarity 33.3%; Pred. No. 2.2e+03;
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 3 LEAQICQOI 11
DB 2 LKSEVSKOI 10

RESULT 14
COG4_CHIOP STANDARD; PRT; 20 AA.
AC P34156;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Collagenolytic protease 23 kDa (EC 3.4.24.7) (Fragment).
OS Chionoecetes opilio (Crab-beetle).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC Brachyura; Eubrachyura; Majoidae; Majidae; Chionoecetes.
OX NCBI_TaxID=41210;
```

```
RN SEQUENCE.
RP TISSUE=Hepatopancreas; PubMed=1663026;
RX MEDLINE=9120073; Vedishcheva Y.V., Strongin A.Y.;
RA Klimova O.A., Vedishcheva Y.V., Strongin A.Y.;
RT "Isolation and characteristics of collagenolytic enzymes from the
RT hepatopancreas of the crab Chionoecetes opilio.";
RL Dokl. Akad. Nauk SSSR 317:482-484(1991).
CC -1- FUNCTION: THIS ENZYME IS A METAL PROTEASE CAPABLE OF DEGRADING
CC THE NATIVE TRIPLE HELIX OF COLLAGEN.
CC -1- CATALYTIC ACTIVITY: Cleaves preferentially one bond in native
CC collagen. Cleavage of the triple helix of collagen at about three-
CC quarters of the length of the molecule from the N-terminus, at
CC 775-Gly-Ile-776 in the alpha-1(I) chain. Cleaves synthetic
CC substrates and alpha-macroglobulins at bonds where PI' is a
CC hydrophobic residue.
CC -1- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A.
DR MEROPS: M12.001;
DR InterPro: IPR000130; Zn_MTpeptidse.
DR PROSITE: PS00142; ZINC_PROTEASE; PARTIAL.
KW Hydrolase; Metalloprotease; Zinc; Collagen degradation.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2108 MW; 2BC7A93D022A97D8 CRC64;

Query Match 20.8%; Score 20; DB 1; Length 20;
Best Local Similarity 50.0%; Pred. No. 2.4e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 IEYVFG 16
DB 15 VPYVFG 20

RESULT 15
FIBB_FELCA STANDARD; PRT; 20 AA.
AC P14469;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE.
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand 19:1789-1791(1965).
CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR InterPro: IPR002181; Fibrinogen_C.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 20 FIBRINOPEPTIDE B.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2328 MW; A829E393B8F627D0 CRC64;

Query Match 20.8%; Score 20; DB 1; Length 20;
Best Local Similarity 75.0%; Pred. No. 2.4e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 IEYV 14
```

Thu Apr 24 08:54:05 2003

us-09-836-073-2.lim25.rsp

Page 6

Db |:|
 2 IDY 5

Search completed: April 23, 2003, 13:43:48
Job time : 5.75169 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:36:43 ; Search time 9.40449 Seconds
(without alignments)
183.999 Million cell updates/sec

Title: US-09-836-073-2
Perfect score: 96
Sequence: 1 AALEAQCQIEYFGDF 18
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 4984

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	33.3	25	2 C57001	endo-1,4-beta-xylanase
2	30	31.2	25	2 A60286	heat-stable serine
3	29	30.2	14	2 PH1679	Ig H chain V-D-J r
4	29	30.2	18	2 PH1368	Ig heavy chain DJ
5	29	30.2	24	2 A45336	cystic fibrosis tr
6	28	29.2	22	2 PH1359	Ig heavy chain DJ
7	27	28.1	12	2 PH0771	T-cell receptor be
8	27	28.1	14	2 PH1598	Ig H chain V-D-J r
9	27	28.1	18	2 PH1629	Ig H chain V-D-J r
10	27	28.1	19	2 A05305	hemoglobin beta-2
11	27	28.1	22	2 PH1325	Ig heavy chain DJ
12	27	28.1	23	2 PH1725	Ig heavy chain V r
13	27	28.1	24	2 PT0258	Ig heavy chain CDR
14	26	28.1	24	2 PH1696	Ig heavy chain V r
15	26	27.1	14	2 PH0755	T-cell receptor be
16	26	27.1	16	2 E53284	T-cell receptor be
17	26	27.1	19	2 B53145	high conductance c
18	26	27.1	20	2 A53592	H-exporting Anpas
19	25.5	26.6	15	2 I78838	fit3 ligand isofor
20	25	26.0	15	2 PH1366	Ig heavy chain DJ
21	25	26.0	17	2 A61211	anantin - Streptom
22	24.5	25.5	14	2 I49514	B144 protein A - m
23	24	25.0	12	2 S57570	T cell receptor V-
24	24	25.0	14	2 S57569	T cell receptor V-
25	24	25.0	14	2 S57638	T cell receptor V-
26	24	25.0	16	1 MTFBBS	melanotropin beta
27	24	25.0	18	1 MTFBFC	melanotropin beta
28	24	25.0	21	2 S78574	protein kinase C 1
29	24	25.0	21	2 PH1730	Ig heavy chain V r

ALIGNMENTS

RESULT 1

C57001
endo-1,4-beta-xylanase (EC 3.2.1.8) 4 - Streptomyces roseiscleroticus (fragment)
C:Species: Streptomyces roseiscleroticus
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 23-May-1997
C:Accession: C57001
R:Grabski, A.C.; Forrester, I.T.; Patel, R.; Jeffries, T.W.
Protein Expr. Purif. 4, 120-129, 1993
A:Title: Characterization and N-terminal amino acid sequences of beta-(1-4)endoxylanase
A:Reference number: A57001; MUID:93229899; PMID:8471845
A:Accession: C57001
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-25 <GRA>
A:Experimental source: strain NRRL B-11019
A:Note: sequence extracted from NCBI backbone (NCBIP:130009)
C:Function:
A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
A:Pathway: xylan degradation
C:Superfamily: Streptomyces endo-1,4-beta-xylanase A; Streptomyces endo-1,4-beta-xy-
C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradati

Query Match 33.3%; Score 32; DB 2; Length 25;
Best Local Similarity 75.0%; Pred. No. 78;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 9 QOIEYFG 16
|||
Db 10 QQSGYFG 17

RESULT 2

A60286
heat-stable serine proteinase (EC 3.4.21.1-) - Thermomonospora fusca (strain YX) (fr
N:Alternate names: YX-proteinase
C:Species: Thermomonospora fusca
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Dec-1996
C:Accession: A60286
R:Kristjansson, M.M.; Kinsella, J.E.
Int. J. Pept. Protein Res. 36, 201-207, 1990
A:Title: Heat stable proteinase from Thermomonospora fusca. Characterization as a s
A:Reference number: A60286; MUID:91107200; PMID:2132918
A:Accession: A60286
A:Molecule type: protein
A:Residues: 1-25 <KRI>
C:Keywords: extracellular protein; hydrolase; serine proteinase

Query Match 31.2%; Score 30; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 13 YVGFDF 18

Db 10 YFNGY 15
|||||:

RESULT 3

PH1626

Ig H chain V-D-J region (clone B-less 118) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999

C:Accession: PH1626

R:Levinson, D.A.; Campos-Torres, J.; Leder, P.

J. Exp. Med. 178, 317-329, 1993

A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice

A:Reference number: PH1580; MUID:93301609; PMID:8315387

A:Accession: PH1626

A:Molecule type: DNA

A:Residues: 1-14 <LEV>

A:Experimental source: bone marrow pre-B lymphocyte

C:Keywords: immunoglobulin

Query Match 30.2%; Score 29; DB 2; Length 14;

Best Local Similarity 40.0%; Pred. NO. 1.4e+02;

Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 CQOIEYVFGD 17

| : : | : |

Db 1 CMRAXYYGD 10

RESULT 4

PH1368

Ig heavy chain DJ region (clone Cl11-112) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: PH1368

R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.

J. Exp. Med. 176, 1577-1581, 1992

A:Title: Predominance of fetal type DJH joining in young children with B precursor lymph

A:Reference number: PH1302; MUID:93094761; PMID:1460419

A:Accession: PH1368

A:Molecule type: DNA

A:Residues: 1-18 <WAS>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 30.2%; Score 29; DB 2; Length 18;

Best Local Similarity 66.7%; Pred. NO. 1.8e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 IEYVFG 16

| : : | : |

Db 5 MEYIYG 10

RESULT 5

A45336

cystic fibrosis transmembrane conductance regulator, CFTR (C-terminal, alternatively sp

C:Species: Homo sapiens (man)

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 20-Aug-1999

C:Accession: A45336

R:Yoshimura, K.; Chu, C.S.; Crystal, R.G.

J. Biol. Chem. 268, 686-690, 1993

A:Title: Alternative splicing of intron 23 of the human cystic fibrosis transmembrane co

minus.

A:Reference number: A45336; MUID:93107081; PMID:7678008

A:Accession: A45336

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-24 <YOS>

A:Cross-references: GB:M96936; NID:gl80293; PIDN:AAB59381.1; PID:gl80295

A:Note: sequence extracted from NCBI backbone (NCBIP:121510)

C:Superfamily: cystic fibrosis transmembrane conductance regulator; ATP-binding cassette

C:Keywords: duplication; glycoprotein; membrane protein; nucleotide binding

Query Match 30.2%; Score 29; DB 2; Length 24;
Best Local Similarity 42.9%; Pred. NO. 2.4e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 AALEAQICQOIEYY 14

| : : | : |

Db 11 AMLECCQOFLQIFY 24

RESULT 6

PH1359

Ig heavy chain DJ region (clone Cl78-121) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: PH1359

R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.

J. Exp. Med. 176, 1577-1581, 1992

A:Title: Predominance of fetal type DJH joining in young children with B precursor l

A:Reference number: PH1302; MUID:93094761; PMID:1460419

A:Accession: PH1359

A:Molecule type: DNA

A:Residues: 1-22 <WAS>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match

29.2%; Score 28; DB 2; Length 22;

Best Local Similarity 44.4%; Pred. NO. 3.2e+02;

Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 8 CQOIEYVFG 16

| : : | : |

Db 6 CYENYYIG 14

RESULT 7

PH0771

T-cell receptor beta chain (PE5.1.1) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999

C:Accession: PH0771

R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.

J. Exp. Med. 174, 1371-1383, 1991

A:Title: T cell receptor genes in a series of class I major histocompatibility compl

allelic exclusion and antigen-specific repertoire

A:Reference number: PH0746; MUID:92078846; PMID:1836010

A:Accession: PH0771

A:Molecule type: mRNA

A:Residues: 1-12 <CAS>

A:Cross-references: EMBL:X60865; NID:g53624; PIDN:CAA43255.1; PID:g53625

A:Experimental source: T lymphocyte

C:Keywords: T-cell receptor

Query Match

28.1%; Score 27; DB 2; Length 12;

Best Local Similarity 71.4%; Pred. NO. 2.5e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYVFG 16

| : : | : |

Db 6 QYEQYFG 12

RESULT 8

PH1598

Ig H chain V-D-J region (wild-type clone 306) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999

C:Accession: PH1598

R:Levinson, D.A.; Campos-Torres, J.; Leder, P.

J. Exp. Med. 178, 317-329, 1993

A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less

A:Reference number: PH1580; MUID:93301609; PMID:8315387

A:Accession: PH1598

A:Molecule type: DNA
 A:Residues: 1-14 <LEV>
 A:Experimental source: bone marrow pre-B lymphocyte
 C:Keywords: immunoglobulin

Query Match 28.1%; Score 27; DB 2; Length 14;
 Best Local Similarity 50.0%; Pred. No. 3e+02;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 8 CQOIEYFGD 17
 I : |||
 Db 1 CAKDGYYGD 10

RESULT 9

PH1629
 Ig H chain V-D-J region (clone B-less 155) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
 C:Accession: PH1629
 R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
 J. Exp. Med. 178, 317-329, 1993
 A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
 A:Reference number: PH1580; MUID:93301609; PMID:8315387
 A:Accession: PH1629
 A:Molecule type: DNA
 A:Residues: 1-18 <LEV>
 A:Experimental source: bone marrow pre-B lymphocyte
 C:Keywords: immunoglobulin

Query Match 28.1%; Score 27; DB 2; Length 18;
 Best Local Similarity 36.4%; Pred. No. 3.9e+02;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 CQOIEYFGDF 18
 I : |||
 Db 1 CVRRDYDGSY 11

RESULT 10

A05305
 hemoglobin beta-2 chain - Indian spiny-tailed lizard (fragment)
 C:Species: Uromastix hardwickii (Indian spiny-tailed lizard)
 C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 01-Aug-1997
 C:Accession: A05305
 R:Nagvi, S.; Zaidi, Z.H.; von Bahr-Lindstrom, H.; Carlquist, M.; Jornvall, H.
 FEBS Lett. 162, 290-295, 1983
 A:Reference number: A91314; MUID:84029159; PMID:6628672
 A:Accession: A05305
 A:Molecule type: protein
 A:Residues: 1-19 <NAQ>
 C:Superfamily: globin; globin homology
 C:Keywords: erythrocyte; oxygen carrier

Query Match 28.1%; Score 27; DB 2; Length 19;
 Best Local Similarity 80.0%; Pred. No. 4.1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 YFGDF 18
 I : |||
 Db 1 FFGDF 5

RESULT 11

PH1325
 Ig heavy chain DJ region (clone C199-121) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C:Accession: PH1325
 R:Wasserman, R.; Gallili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
 J. Exp. Med. 176, 1577-1581, 1992
 A:Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma
 A:Reference number: PH1302; MUID:93094761; PMID:1460419

A:Accession: PH1325
 A:Molecule type: DNA
 A:Residues: 1-22 <WAS>
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

Query Match 28.1%; Score 27; DB 2; Length 22;
 Best Local Similarity 66.7%; Pred. No. 4.8e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 IEYFYG 16
 I : |||
 Db 9 IHYYG 14

RESULT 12

PH1725
 Ig heavy chain V region (clone GCC-5) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
 C:Accession: PH1725
 R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
 J. Exp. Med. 178, 295-307, 1993
 A:Title: Antigen-driven B cell differentiation in vivo.
 A:Reference number: PH1675; MUID:93301607; PMID:8315385
 A:Accession: PH1725
 A:Molecule type: mRNA
 A:Residues: 1-23 <MCH>
 A:Experimental source: B cell
 A:Note: the authors translated the codon ACA for residue 13 as Ala
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

Query Match 28.1%; Score 27; DB 2; Length 23;
 Best Local Similarity 33.3%; Pred. No. 5e+02;
 Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 10 QIEYFGDF 18
 I : |||
 Db 12 ETRYGYSY 20

RESULT 13

PT0258
 Ig heavy chain CDR3 region (clone 2-118B) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 30-May-1997
 C:Accession: PT0258
 R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
 J. Exp. Med. 173, 395-407, 1991
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity
 A:Reference number: PT0222; MUID:91108337; PMID:1899102
 A:Accession: PT0258
 A:Molecule type: DNA
 A:Residues: 1-24 <YAM>
 A:Experimental source: B lymphocyte
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

Query Match 28.1%; Score 27; DB 2; Length 24;
 Best Local Similarity 33.3%; Pred. No. 5.3e+02;
 Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 AQICQOIEYFG 16
 I : |||
 Db 5 AEYSSLYYVG 16

RESULT 14

PH1696
 Ig heavy chain V region (clone NP-7-9) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999

C;Accession: PH1696
R;McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A;Title: Antigen-driven B cell differentiation in vivo.
A;Reference number: PH1675; MUID:93301607; PMID:8315385
A;Accession: PH1696
A;Molecule type: mRNA
A;Residues: 1-24 <MCH>
A;Experimental source: B cell
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 28.1%; Score 27; DB 2; Length 24;
Best Local Similarity 42.9%; Pred. No. 5.3e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 10 QIEYYFG 16
::|||
Db 12 EVAYYG 18

RESULT 15

PH0755

T-cell receptor beta chain (QAll.3.2) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
C;Accession: PH0755
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-I allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0755
A;Molecule type: mRNA
A;Residues: 1-14 <CAS>
A;Cross-references: EMBL:X60849; NID:g53876; PIDN:CAA43240.1; PID:g53877
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match 27.1%; Score 26; DB 2; Length 14;
Best Local Similarity 62.5%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 QIEYYFG 16
|||
Db 7 QPYEQYFG 14

Search completed: April 23, 2003, 13:48:53
Job time : 10.4045 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:43:09 ; Search time 9.91011 seconds
(without alignments)
145.542 Million cell updates/sec

Title: US-09-836-073-2
Perfect score: 96
Sequence: 1 AALEAQCQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 61792

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96	100.0	18	US-09-836-073-2	Sequence 2, Appli
2	87	90.6	18	US-09-836-073-1	Sequence 1, Appli
3	87	90.6	18	US-09-836-073-14	Sequence 14, Appl
4	83	86.5	17	US-09-836-073-13	Sequence 13, Appl
5	81	84.4	18	US-09-836-073-9	Sequence 9, Appli
6	79	82.3	18	US-09-836-073-11	Sequence 11, Appl
7	79	82.3	18	US-09-836-073-12	Sequence 12, Appl
8	78	81.2	18	US-09-836-073-10	Sequence 10, Appl
9	76.5	79.7	19	US-09-836-073-16	Sequence 16, Appl
10	75	78.1	18	US-09-836-073-3	Sequence 3, Appli
11	75	78.1	18	US-09-836-073-4	Sequence 4, Appli
12	74	77.1	18	US-09-836-073-15	Sequence 15, Appl
13	72	75.0	18	US-08-836-073-7	Sequence 7, Appli
14	71	74.0	18	US-09-836-073-8	Sequence 8, Appli
15	63	65.6	18	US-09-836-073-5	Sequence 5, Appli
16	58	60.4	16	US-09-836-073-19	Sequence 19, Appl
17	56	58.3	18	US-09-836-073-6	Sequence 6, Appli
18	45	46.9	18	US-09-836-073-17	Sequence 17, Appl
19	43.5	45.3	18	US-09-836-073-18	Sequence 18, Appl

20	30	31.2	19	9	US-10-153-159-6	Sequence 6, Appli
21	30	31.2	19	9	US-10-153-159-51	Sequence 51, Appl
22	30	31.2	19	9	US-10-153-159-53	Sequence 53, Appl
23	30	31.2	19	9	US-10-153-159-54	Sequence 54, Appl
24	30	31.2	19	9	US-10-153-176-6	Sequence 6, Appli
25	30	31.2	19	9	US-10-153-176-51	Sequence 51, Appl
26	30	31.2	19	9	US-10-153-176-53	Sequence 53, Appl
27	30	31.2	19	9	US-10-153-176-54	Sequence 54, Appl
28	30	31.2	21	10	US-09-864-761-33741	Sequence 33741, A
29	30	30.2	13	9	US-09-826-290-280	Sequence 280, App
30	29	30.2	13	9	US-09-791-389-192	Sequence 192, App
31	29	30.2	13	9	US-09-791-393-192	Sequence 192, App
32	29	30.2	13	10	US-09-791-378-452	Sequence 452, App
33	29	30.2	19	9	US-10-102-806-802	Sequence 802, App
34	29	30.2	25	9	US-10-042-431-51	Sequence 51, Appl
35	29	30.2	25	9	US-09-759-1308-421	Sequence 421, App
36	28	29.2	15	10	US-09-073-009-51	Sequence 51, Appl
37	28	29.2	15	10	US-09-023-588-51	Sequence 51, Appl
38	28	29.2	15	10	US-09-793-306-51	Sequence 51, Appl
39	28	29.2	16	9	US-10-038-612-94	Sequence 94, Appl
40	28	29.2	18	9	US-10-084-813-139	Sequence 139, App
41	28	29.2	18	9	US-10-084-813-140	Sequence 140, App
42	28	29.2	18	9	US-10-084-813-141	Sequence 141, App
43	28	29.2	19	9	US-10-153-159-49	Sequence 49, Appl
44	28	29.2	19	9	US-10-153-159-59	Sequence 59, Appl
45	28	29.2	19	9	US-10-153-176-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baldya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match 100.0%; Score 96; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.3e+08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AALEAQCQIEYFGDF 18
|||||
Db 1 AALEAQCQIEYFGDF 18

RESULT 2
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baldya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match 90.68; Score 87; DB 9; Length 18;
Best Local Similarity 88.98; Pred. No. 3.3e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIEYYFGDF 18
||||| |||||
Db 1 AALEAKICHQIEYYFGDF 18

RESULT 3

US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14

Query Match 90.68; Score 87; DB 9; Length 18;
Best Local Similarity 88.98; Pred. No. 3.3e-07;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIEYYFGDF 18
||||| |||||
Db 1 AALEAKICHQIEYYFGDF 18

RESULT 4

US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match 86.58; Score 83; DB 9; Length 17;
Best Local Similarity 88.28; Pred. No. 1.3e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 AALEAQCQIEYYFGDF 18
||||| |||||
Db 1 AALEAKICHQIEYYFGDF 17

RESULT 5

US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match 84.48; Score 81; DB 9; Length 18;
Best Local Similarity 88.28; Pred. No. 2.8e-06;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AALEAQCQIEYYFGD 17
||||| |||||
Db 1 AALEAKICHQIEYYFGD 17

RESULT 6

US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Balidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match 82.38; Score 79; DB 9; Length 18;
Best Local Similarity 83.38; Pred. No. 5.7e-06;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALEAQCQIEYYFGDF 18
||||| |||||
Db 1 AALEAKICHQIEYYFGDF 18

RESULT 7

US-09-836-073-12

; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match 82.3%; Score 79; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 5.7e-06;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALEAQICQOIYYFGDF 18
|||||:|||||
Db 1 AALEAKICHQIEYQGF 18

RESULT 8
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match 81.2%; Score 78; DB 9; Length 18;
Best Local Similarity 83.3%; Pred. No. 8.1e-06;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALEAQICQOIYYFGDF 18
|||||:|||||
Db 1 AALEAKICHQIEYQGF 18

RESULT 9
US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match 79.7%; Score 76.5; DB 9; Length 19;
Best Local Similarity 84.2%; Pred. No. 1.5e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALEAQICQOIYYFGDF 18
|||||:|||||
Db 1 AALEAKICHQIEYFGDF 19

RESULT 10
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match 78.1%; Score 75; DB 9; Length 18;
Best Local Similarity 72.2%; Pred. No. 2.4e-05;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 AALEAQICQOIYYFGDF 18
|||||:|||||
Db 1 AALEAKICHQIEYQGF 18

RESULT 11
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match 78.1%; Score 75; DB 9; Length 18;
Best Local Similarity 86.7%; Pred. No. 2.4e-05;

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:38:28 ; Search time 8.49438 Seconds
(without alignments)
62.349 Million cell updates/sec

Title: US-09-836-073-2
Perfect score: 96
Sequence: 1 AALEAQICQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 147762

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	87	90.6	18	4	US-09-316-630-3
2	87	90.6	18	4	US-09-316-630-4
3	35	36.5	14	2	US-08-433-133-94
4	32	33.3	25	1	US-08-453-289-4
5	32	33.3	25	2	US-08-574-086-4
6	30	31.2	11	2	US-08-618-696-7
7	30	31.2	11	3	US-09-033-753-7
8	30	31.2	15	1	US-07-603-675-3
9	30	31.2	24	4	US-07-963-329A-73
10	30	31.2	24	5	PCT-US92-09443A-73
11	29	30.2	10	2	US-08-618-696-11
12	29	30.2	10	3	US-09-033-753-11
13	29	30.2	11	2	US-08-618-696-20
14	29	30.2	11	3	US-09-033-753-20
15	29	30.2	12	4	US-09-298-924-55
16	29	30.2	22	6	5281520-43
17	28	29.2	10	2	US-08-618-696-2
18	28	29.2	10	2	US-08-618-696-10
19	28	29.2	10	3	US-09-033-753-2
20	28	29.2	10	3	US-09-033-753-10
21	28	29.2	11	2	US-08-618-696-1
22	28	29.2	11	2	US-08-618-696-6
23	28	29.2	11	2	US-08-618-696-16
24	28	29.2	11	2	US-08-618-696-19
25	28	29.2	11	3	US-09-033-753-1
26	28	29.2	11	3	US-09-033-753-6
27	28	29.2	11	3	US-09-033-753-16

28	28	29.2	11	3	US-09-033-753-19	Sequence 19, Appl
29	28	29.2	12	2	US-08-968-676-165	Sequence 165, App
30	28	29.2	13	2	US-08-480-190-47	Sequence 47, Appl
31	28	29.2	13	2	US-08-488-379-47	Sequence 47, Appl
32	28	29.2	13	5	PCT-US93-07545-47	Sequence 47, Appl
33	28	29.2	13	5	PCT-US94-10257A-27	Sequence 27, Appl
34	28	29.2	14	2	US-08-637-759B-113	Sequence 113, App
35	28	29.2	14	3	US-08-871-355A-113	Sequence 113, App
36	28	29.2	14	4	US-09-307-143-10	Sequence 10, Appl
37	28	29.2	14	4	US-09-201-945-113	Sequence 113, Appl
38	28	29.2	15	1	US-07-715-934-16	Sequence 16, Appl
39	28	29.2	15	1	US-08-213-124-24	Sequence 24, Appl
40	28	29.2	16	1	US-08-305-871A-13	Sequence 13, Appl
41	28	29.2	16	4	US-08-788-822A-13	Sequence 13, Appl
42	27.5	28.6	12	1	US-08-044-631-1	Sequence 1, Appl
43	27.5	28.6	12	1	US-08-161-907-5	Sequence 5, Appl
44	27.5	28.6	12	1	US-08-480-604A-31	Sequence 31, Appl
45	27.5	28.6	12	1	US-08-456-847-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316.630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match 90.6%; Score 87; DB 4; Length 18;
Best Local Similarity 88.9%; Pred. No. 2.7e-08;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 AALEAQICQIEYFGDF 18
| | | | | | | | | | | | | | | | | | | | | |
Db 1 AALEAKICHQIEYFGDF 18

RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316.630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

Query Match	33.3%	Score 32;	DB 1;	Length 25;
Best Local Similarity	75.0%	Pred. No. 54;		

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 QQIEYFG 16
|||
Db 10 QQSGYFG 17

RESULT 5

US-08-574-086-4
; Sequence 4, Application US/08574086
; Patent No. 5834301
; GENERAL INFORMATION:
; APPLICANT: Jeffries, Thomas W
; APPLICANT: Grabski, Anthony C
; APPLICANT: Patel, Rajesh N
; APPLICANT: Elegir, Graziano
; APPLICANT: Szakacs, George
; TITLE OF INVENTION: Method of Removing Color from Kraft Wood
; TITLE OF INVENTION: Pulp
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janet I. Stockhausen
; STREET: One Gifford Pinchot Drive
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53705

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,086
; FILING DATE: 18-DEC-1995

CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,289
; FILING DATE:
; FILING DATE:
; APPLICATION NUMBER: US/08/257,965
; FILING DATE:
; APPLICATION NUMBER: US 07/857,060
; FILING DATE: 25-MAY-1992

ATTORNEY/AGENT INFORMATION:
; NAME: Stockhausen, Janet I
; REGISTRATION NUMBER: 34,256
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-231-9504
; TELEFAX: 608-231-9508
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptomyces roseisclerotiscus
; STRAIN: NRRLB-11019

US-08-574-086-4
Query Match 33.3%; Score 32; DB 2; Length 25;
Best Local Similarity 75.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 QQIEYFG 16
|||
Db 10 QQSGYFG 17

RESULT 6

US-08-618-696-7
; Sequence 7, Application US/08618696
; Patent No. 5861475
; GENERAL INFORMATION:
; APPLICANT: COOPER, Jr., J. ALLEN D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210

COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,696
; FILING DATE: 20-MAR-1996

CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/995,269
; FILING DATE: 12/21/92
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.

REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UOAB:002/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-320-7200
; TELEFAX: 512-474-7577
; TELEX: NOT APPLICABLE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-618-696-7
Query Match 31.2%; Score 30; DB 2; Length 11;
Best Local Similarity 44.4%; Pred. No. 45;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 9 QQIEYFGD 17
::|::|
Db 1 REASYFGD 9

RESULT 7

US-09-033-753-7
; Sequence 7, Application US/09033753
; Patent No. 6017883
; GENERAL INFORMATION:
; APPLICANT: COOPER, Jr., J. ALLEN D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210

COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1


```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/033,753
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/618,696
; FILING DATE: 20-MAR-1996
; APPLICATION NUMBER: 07/995,269
; FILING DATE: 12/21/92
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UOAB:002/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-320-7200
; TELEFAX: 512-474-7577
; TELEX: NOT APPLICABLE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-033-753-7

```

```

Query Match 31.28; Score 30; DB 3; Length 11;
Best Local Similarity 44.4%; Pred. No. 45;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 9 QQIEYFGD 17
   : : : : :
DB 1 REASYFGD 9

```

```

RESULT 8
US-07-603-675-3
; Sequence 3, Application US/07603675
; Patent No. 5416006
; GENERAL INFORMATION:
; APPLICANT: Blas1, Francesco
; APPLICANT: Stoppelli, Maria P
; APPLICANT: Mastronicola, Maria R
; APPLICANT: Wellinder, Karen G
; APPLICANT: Correas, Isabel
; TITLE OF INVENTION: MODIFICATION OF PLASMINOGEN ACTIVATORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOPER & DUNHAM
; STREET: 30 ROCKEFELLER PLAZA
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/603,675
; FILING DATE: 19911218
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/DK90/00096
; FILING DATE: 11-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 38154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 644-0525
; TELEX: (212) 422523 COOP UI

```

```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: N
; FRAGMENT TYPE: internal
US-07-603-675-3

```

```

Query Match 31.28; Score 30; DB 1; Length 15;
Best Local Similarity 71.4%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 8 CQIEY 14
   : : : : :
DB 2 CQPHY 8

```

```

RESULT 9
US-07-963-329A-73
; Sequence 73, Application US/07963329A
; Patent No. 6310040
; GENERAL INFORMATION:
; APPLICANT: Bozycko-Coyne, Donna
; APPLICANT: Neff, Nicola
; APPLICANT: Lewis, Michael E.
; APPLICANT: Iqbal, Mohamed
; TITLE OF INVENTION: TREATING RETINAL NEURONAL DISORDERS
; TITLE OF INVENTION: BY THE APPLICATION OF INSULIN-LIKE
; TITLE OF INVENTION: GROWTH FACTORS AND ANALOGS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/963,329A
; FILING DATE: 19921015
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/790,690
; FILING DATE: NO. 6310040ember 8, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 02655/012002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-07-963-329A-73

```

```

Query Match 31.28; Score 30; DB 4; Length 24;
Best Local Similarity 30.8%; Pred. No. 1.1e+02;
Matches 4; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 5 AQICQIEYFGD 17
   : : : : :

```

Db 2 AELVDALQFYSGD 14

RESULT 10

PCT-US92-09443A-73
; Sequence 73, Application PC/TUS9209443A

; GENERAL INFORMATION:

; APPLICANT: Bozyczko-Coyne, Donna

; APPLICANT: Neff, Nicola

; APPLICANT: Lewis, Michael E.

; APPLICANT: Iqbal, Mohamed

; TITLE OF INVENTION: TREATING RETINAL NEURONAL

; TITLE OF INVENTION: DISORDERS BY THE APPLICATION OF

; TITLE OF INVENTION: INSULIN-LIKE GROWTH FACTORS AND

; TITLE OF INVENTION: ANALOGS

; NUMBER OF SEQUENCES: 79

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; COMPUTER: IBM PS/2 Model 502 or 55SX

; OPERATING SYSTEM: MS-DOS (Version 5.0)

; SOFTWARE: WordPerfect (Version 5.1)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/09443A

; FILING DATE: 19921103

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/790,690

; FILING DATE: November 8, 1991

; APPLICATION NUMBER: 07/963,329

; FILING DATE: October 15, 1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Clark, Paul T.

; REGISTRATION NUMBER: 30,162

; REFERENCE/DOCKET NUMBER: 02655/012WO2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 542-5070

; TELEFAX: (617) 542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 73:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 24

; TYPE: AMINO ACID

; STRANDEDNESS: N/A

; TOPOLOGY: N/A

PCT-US92-09443A-73

Query Match 31.2%; Score 30; DB 5; Length 24;

Best Local Similarity 30.8%; Pred. No. 1.1e+02;

Matches 4; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 AQICQIEYYFGD 17

; : : : : | |

Db 2 AELVDALQFYSGD 14

RESULT 11

US-08-618-696-11

; Sequence 11, Application US/08618696

; Patent No. 5861475

; GENERAL INFORMATION:

; APPLICANT: COOPER, Jr., J. ALLEN D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; TITLE OF INVENTION: INHIBITION OF PHAGOCYTES

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ARNOLD, WHITE & DURKEE

STREET: P.O. BOX 4433

CITY: HOUSTON

STATE: TEXAS

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/618,696

FILING DATE: 20-MAR-1996

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/995,269

FILING DATE: 12/21/92

ATTORNEY/AGENT INFORMATION:

NAME: PARKER, DAVID L.

REGISTRATION NUMBER: 32,165

REFERENCE/DOCKET NUMBER: UOAB:002/PAR

TELECOMMUNICATION INFORMATION:

TELEPHONE: 512-320-7200

TELEFAX: 512-474-7577

TELEX: NOT APPLICABLE

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acid residues

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-618-696-11

Query Match 30.2%; Score 29; DB 2; Length 10;

Best Local Similarity 50.0%; Pred. No. 59;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYYFGD 17

; : : : : | | | |

Db 1 EASYFFGD 8

RESULT 12

US-09-033-753-11

; Sequence 11, Application US/09033753

; Patent No. 6017883

; GENERAL INFORMATION:

; APPLICANT: COOPER, Jr., J. ALLEN D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE

; TITLE OF INVENTION: INHIBITION OF PHAGOCYTES

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

ADDRESSEE: ARNOLD, WHITE & DURKEE

STREET: P.O. BOX 4433

CITY: HOUSTON

STATE: TEXAS

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/033,753

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/618,696

FILING DATE: 20-MAR-1996

APPLICATION NUMBER: 07/995,269

FILING DATE: 12/21/92

ATTORNEY/AGENT INFORMATION:

NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UOAB:002/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577
TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-033-753-11

Query Match 30.2%; Score 29; DB 3; Length 10;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYFGD 17
Db 1 EASYFGD 8

RESULT 13
US-08-618-696-20
Sequence 20, Application US/08618696
Patent No. 5861475

GENERAL INFORMATION:
APPLICANT: COOPER, JR., J. ALLEN D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSER: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/618,696
FILING DATE: 20-MAR-1996
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/995,269
FILING DATE: 12/21/92
ATTORNEY/AGENT INFORMATION:
NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UOAB:002/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577

TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-618-696-20

Query Match 30.2%; Score 29; DB 2; Length 11;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYFGD 17

Db 2 EASYFGD 9

RESULT 14

US-09-033-753-20
Sequence 20, Application US/09033753
Patent No. 6017883
GENERAL INFORMATION:
APPLICANT: COOPER, JR., J. ALLEN D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TITLE OF INVENTION: INHIBITION OF PHAGOCYTES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARNOLD, WHITE & DURKEE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,753
FILING DATE:

CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/618,696
FILING DATE: 20-MAR-1996
APPLICATION NUMBER: 07/995,269
FILING DATE: 12/21/92
ATTORNEY/AGENT INFORMATION:

NAME: PARKER, DAVID L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UOAB:002/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512-320-7200
TELEFAX: 512-474-7577

TELEX: NOT APPLICABLE
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-033-753-20

Query Match 30.2%; Score 29; DB 3; Length 11;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYFGD 17
Db 2 EASYFGD 9

RESULT 15

US-09-298-924-55
Sequence 55, Application US/09298924
Patent No. 6391595
GENERAL INFORMATION:

APPLICANT: KATO, Masaru
MIURA, Yutaka
KETTOKU, Masako
IWAMATSU, Akihiro
KOBAYASHI, Kazuo
KOMEDA, Toshihiro

TITLE OF INVENTION: NOVEL TRANSFERASE AND AMYLASE, PROCESS
FOR PRODUCING THE ENZYMES, USE THEREOF, AND GENE CODING
FOR THE SAME

```

;
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/298,924
; FILING DATE: 26-Apr-1999
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/750,569
; FILING DATE: <Unknown>
; APPLICATION NUMBER: JP 7-120673
; FILING DATE: 21-APR-1995
; APPLICATION NUMBER: JP 6-311185
; FILING DATE: 21-NOV-1994
; APPLICATION NUMBER: JP 6-286917
; FILING DATE: 21-NOV-1994
; APPLICATION NUMBER: JP 6-290394
; FILING DATE: 31-OCT-1994
; APPLICATION NUMBER: JP 6-194223
; FILING DATE: 18-AUG-1994
; APPLICATION NUMBER: JP 6-133354
; FILING DATE: 16-JUN-1994
;
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 49441/110
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
;
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 55:
;
; US-09-298-924-55

```

```

Query Match 30.2%; Score 29; DB 4; Length 14;
Best Local Similarity 57.1%; Pred. No. 87;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 12 EYFEGDF 18
Db 1 DYYIQDF 7

```

```

Search completed: April 23, 2003, 13:50:24
Job time : 9.49438 secs

```

CC IgE134 residues (R). (C) is useful for inhibiting the binding of IgE to
 CC high affinity IgE receptor (FcεRI). Peptides of the formula given
 CC in the specification are useful for inhibiting the binding of an IgE to
 CC high affinity IgE receptor. The peptide is useful for selecting a
 CC molecule which blocks the interaction of IgE with high affinity IgE
 CC receptor. The peptide is also useful for inhibiting the activation of
 CC high affinity IgE receptor. The peptide is useful for treating an IgE-
 CC mediated disease or disorder in a host. (C) is useful in research,
 CC diagnostic, therapeutic and prophylactic methods. The peptide is also
 CC useful for inhibiting IgE-mediated or associated processes such as IgE-
 CC dependent activation and degranulation of mast cells and basophils, as
 CC well as consequent release of inflammatory mediators such as histamine.
 CC (C) is useful for treating allergic rhinitis, asthma (e.g. allergic
 CC asthma), atopic dermatitis, urticaria-angioedema, parasitic infection,
 CC IgE myeloma, immune-related disorders, inflammatory disorders, diabetes
 CC mellitus, IgE-mediated gastrointestinal inflammatory disease, immune
 CC rejection of grafts, reperfusion injury, stroke, myocardial infarction,
 CC atherosclerosis, acute lung injury, haemorrhagic shock, burn, septic
 CC shock, acute tubular necrosis, endometriosis, degenerative joint disease
 CC and pancreatitis. The present sequence is a peptide of the invention
 CC expressed from a phage display library.

XX SQ Sequence 21 AA;

Query Match 32.3%; Score 31; DB 23; Length 21;
 Best Local Similarity 40.0%; Pred. No. 3e+02;
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICQIIEYFG 16
 :|::|||
 Db 2 VCPRLCYWFG 11

RESULT 15

AAU00592
 ID AAU00592 standard; Peptide; 23 AA.

AC AAU00592;

XX 18-MAY-2001 (first entry)

XX VH ligand-binding domain variant CDR3 region #8.

XX Human; monoclonal antibody A6; VH ligand-binding domain; variable region;
 KW parental ligand binding molecule; PLBM; framework region; FR; CDR;
 KW combinatorial library; diagnosis; therapy; surface protein; pesticide;
 KW cell-receptor; carbohydrate; lipid; antibiotic; steroid; hormone; dye;
 KW herbicide; polymer; immunoglobulin; CDR3 region; M2 immunoglobulin;
 KW panning.

XX Homo sapiens.

XX WO200118058-A2.

XX 15-MAR-2001.

XX 07-SEP-2000; 2000WO-CA01027.

XX 07-SEP-1999; 99CA-2282179.

XX 04-NOV-1999; 99US-0163546.

XX (NOVO-) NOVOPHARM BIOTECH INC.

XX Kaplan H, Entwistle J, Tanha J, Narang S, Dan M;

XX WPI; 2001-235191/24.

XX Combinatorial libraries including phage display library comprises
 PT variants of immunoglobulin VH fragments which comprises the framework
 PT regions of wild-type or modified immunoglobulin VH domain of human A6
 PT antibody

XX Disclosure; page 25; 133pp; English.

XX The sequence represents a variant CDR3 region of a parental VH
 CC ligand-binding domain of human monoclonal antibody A6, isolated by
 CC panning against M2 immunoglobulin. Variants of a parental ligand
 CC binding molecule (PLBM) comprise an immunoglobulin (Ig) VH fragment
 CC including the framework (FR) regions of the wild-type/modified Ig VH
 CC domain of human monoclonal antibody A6. The variant differs from PLBM at
 CC residues constituting part of one of the CDRs of PLBM. A library
 CC expressing binding domains can be created by cloning a parental DNA
 CC sequence that encodes a parental domain, in order to produce parental
 CC clones, and replacing a variable region of the clones with a variant DNA
 CC sequence. A library of genetic packages can then be generated, each
 CC package having a surface and a surface protein binding domain expressed
 CC by the variant DNA. Combinatorial libraries are useful for generating
 CC forms of recombinant antibody fragments that specifically bind target
 CC ligands including natural and synthetic molecules and macromolecules to
 CC be used in diagnostic and/or therapeutic processes. Examples include
 CC enzymes, cell-receptors, carbohydrates, lipids, and organic-based and
 CC inorganic-based molecules such as antibiotics, steroids, hormones,
 CC pesticides, herbicides, dyes and polymers.

XX SQ Sequence 23 AA;

Query Match 32.3%; Score 31; DB 22; Length 23;
 Best Local Similarity 41.2%; Pred. No. 3.3e+02;
 Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 ALEAQICQIIEYFGDF 18
 |||||:
 Db 5 ALSPPQACMTKERYKDF 21

Search completed: April 23, 2003, 13:42:55
 Job time : 28.2022 secs

XX PI Mack M, Schloendorff D, Spring M;
 XX WPI; 2002-362240/39.
 XX Use of an antibody and/or chemokine construct that binds to a chemokine
 PT receptor, for eliminating cells latently infected with primate
 PT immunodeficiency virus, or treating, preventing and alleviating immune
 PT disorders -
 XX Disclosure; Page 117; 117pp; English.
 XX The invention relates to the use of an antibody and/or chemokine
 CC construct that binds a chemokine receptor for preparing a pharmaceutical
 CC composition for eliminating cells latently infected with a primate
 CC immunodeficiency virus (e.g. human immunodeficiency virus (HIV-1)). They
 CC are used in gene therapy and as vaccines. The antibody and/or chemokine
 CC construct is also used for preparing a pharmaceutical composition for
 CC treating, preventing and/or alleviating immunological disorders
 CC including autoimmune diseases (e.g. multiple sclerosis, type I diabetes
 CC and rheumatoid arthritis), allergic diseases, skin diseases (e.g. skin
 CC inflammation, atopic dermatitis and psoriasis), inflammatory diseases
 CC such as inflammatory joint diseases (chronic arthritis), inflammatory
 CC renal diseases and inflammatory bowel diseases and graft versus host
 CC and transplant rejections. The present sequence is murine MC-1 antibody
 CC heavy chain variable domain (VH) complementarity determining region 3
 CC (CDR3) peptide.
 XX Sequence 10 AA;
 SQ Query Match 32.3%; Score 31; DB 23; Length 10;
 Best Local Similarity 71.4%; Pred. NO. 1.3e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 12 EYFGDF 18
 |||||
 Db 2 EYVYGIF 8
 RESULT 13
 AAP90453
 ID AAP90453 standard; protein; 19 AA.
 XX AAP90453;
 AC AAP90453;
 XX 10-NOV-1989 (first entry)
 DT Epitope recognised by anti-amphiregulin antibody.
 XX Amphiregulin; growth regulator; neoplastic disease; bone resorption;
 DE immune response; arachidonic acid cascade.
 KW Breast carcinoma cell line MCF-7.
 XX DE3902157-A.
 OS 27-JUL-1989.
 PN 25-JAN-1989; 89DE-3902157.
 PD 25-JAN-1988; 88US-0148327.
 XX 15-APR-1988; 88US-0181884.
 PR 17-JAN-1989; 89US-0297816.
 XX (ONCO) ONCOGEN.
 PA Shoyab M, McDonald VL, Plowman G, Bradley JG;
 PI WPI; 1989-221730/31.
 XX New growth regulating proteins designated amphiregulin - inhibiting
 PT tumours but stimulating some normal cells, also new precursor protein,
 PT encoding nucleic acid sequences and antibodies.

XX Claim 48; page 41; 78pp; German.
 XX The sequence shows an epitope recognised by an anti-amphiregulin
 CC antibody. See also AAP90448-9 and AAP90372-6.
 CC Sequence 19 AA;
 SQ Query Match 32.3%; Score 31; DB 10; Length 19;
 Best Local Similarity 42.9%; Pred. NO. 2.7e+02;
 Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 4 EAQICQIEYFFGD 17
 |||||
 Db 1 EAVTCRCQEQEYFGE 14
 RESULT 14
 ABG66356
 ID ABG66356 standard; Peptide; 21 AA.
 XX ABG66356;
 AC ABG66356;
 XX 30-AUG-2002 (first entry)
 DT IgE Fc epsilon RI binding peptide IGE120 #20.
 DE IgE receptor; immunoglobulin; Fc epsilon RI; antagonist; phage display;
 KW protein co-ordinate data; IgE-mediated disease; allergic rhinitis;
 KW asthma; allergic asthma; atopic dermatitis; urticaria-angioedema;
 KW parasitic infection; IgE myeloma; immune-related disorder;
 KW inflammatory disorder; diabetes mellitus; reperfusion injury; stroke;
 KW IgE-mediated gastrointestinal inflammatory disease; burn;
 KW immune rejection of graft; myocardial infarction; atherosclerosis;
 KW acute lung injury; haemorrhagic shock; septic shock;
 KW acute tubular necrosis; endometriosis; degenerative joint disease;
 KW pancreatitis.
 XX Synthetic.
 OS WO200226781-A2.
 PN 04-APR-2002.
 PD 26-SEP-2001; 2001WO-US30289.
 PF 26-SEP-2000; 2000US-235353P.
 XX 23-MAR-2001; 2001US-278540P.
 PR (GETH) GENENTECH INC.
 XX Lowman HB, Reynolds ME, Nakamura GR, Starovasnik MA;
 PI WPI; 2002-444016/47.
 DR A peptide useful for treating a IgE-mediated disease or disorder in a
 PT host e.g. allergic rhinitis, asthma, which competes with immunoglobulin
 PT E for binding to high affinity IgE receptor in an in vitro assay -
 XX Example 8; Fig 2; 328pp; English.
 PS The invention relates to a peptide which competes with immunoglobulin
 CC (Ig) E 134 comprising a sequence (S1), for binding the high affinity IgE
 CC receptor (Fc epsilon RI) in an in vitro assay and having a formula given in
 CC the specification. Also included are a fusion protein comprising the
 CC peptide, a pharmaceutical composition (C) comprising the peptide,
 CC designing a compound that mimics the three-dimensional surface
 CC structure of the peptide, a compound with a solvent accessible surface
 CC that mimics the solvent accessible surface defined by the side chains of
 CC residues (R) Pro4, Phe6, Prol6, Cys3, Cys7, Cys15 and Cys19 of IGE134,
 CC a peptide with structural coordinates as given in the specification,
 CC selecting a peptide mimetic which binds to Fc epsilon RI and blocks
 CC binding of IgE and a peptide mimetic which mimics the coordinates of

OS Streptomyces roseiscleroticus.
 PN US5834301-A.
 XX
 PD 10-NOV-1998.
 XX
 PF 18-DEC-1995; 95US-0574086.
 XX
 PR 08-JUN-1994; 94US-0257965.
 PR 25-MAR-1992; 92US-0857060.
 PR 30-MAY-1995; 95US-0453289.
 PR 18-DEC-1995; 95US-0574086.
 XX
 PA (USDA) US SEC OF AGRIC.
 XX
 PI Elegir G, Grabski AC, Jeffries TW, Patel RN, Szakacs G;
 XX
 DR WPI; 1999-008727/01.
 XX
 PT Removal of colour from wood pulp by treatment with specific
 PT xylanases - from Streptomyces strains NRRL 18982 or 18984 and then
 PT extracting the chromophores to reduce bleach demand in subsequent
 PT stages
 XX
 PS Example 1; Fig 6; 45pp; English.
 XX
 CC AAW80364-67 represent the N-terminal sequences of xylanases of
 CC S. roseiscleroticus strain NRRLB-11019. The xylanases are used in a
 CC method for removing colour from wood pulp. The method comprises
 CC treating with a xylanase, then extraction to remove chromophores.
 CC By removing chromophores (formed during the kraft cooking process),
 CC the method reduces the amount of bleach required in subsequent stages.
 CC The method is applied to kraft pulp from softwood or hardwood, or to
 CC pulp from secondary sources.
 XX
 SQ Sequence 25 AA;
 Query Match 33.3%; Score 32; DB 20; Length 25;
 Best Local Similarity 75.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 9 QQIEYFYG 16
 DB |||||
 10 QQSGYFYG 17
 RESULT 11
 AAR43395
 ID AAR43395 standard; peptide; 8 AA.
 XX
 AC AAR43395;
 XX
 DT 12-MAY-1994 (first entry)
 XX
 DE La/SSB epitope 24.
 XX
 KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9321223-A.
 XX
 PD 28-OCT-1993.
 XX
 PF 13-APR-1993; 93WO-US03484.
 XX
 PR 13-APR-1992; 92US-0867819.
 XX
 PA (OKLA) UNIV OKLAHOMA STATE.
 XX

PI Harley JB;
 XX
 DR WPI; 1993-351658/44.
 XX
 PT New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribo:nucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus
 XX
 PS Claim 1; Page 30; 43pp; English.
 XX
 CC The sequences given in AAR43391-562 are linear epitopes which are
 CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polynuridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.
 XX
 SQ Sequence 8 AA;
 Query Match 32.3%; Score 31; DB 14; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 14 YFGDF 18
 DB |||||
 1 YFGDF 5
 RESULT 12
 AAE22202
 ID AAE22202 standard; peptide; 10 AA.
 XX
 AC AAE22202;
 XX
 DT 25-JUL-2002 (first entry)
 XX
 DE Murine MC-1 antibody heavy chain variable domain (VH) CDR3 peptide.
 XX
 KW Chemokine construct; human immunodeficiency virus 1; allergic disease;
 KW skin disease; immunological disorder; autoimmune disease; psoriasis;
 KW multiple sclerosis; rheumatoid arthritis; inflammatory bowel disease;
 KW diabetes; skin inflammation; atopic dermatitis; inflammatory disease;
 KW inflammatory renal disease; HIV-1; transplant rejection; murine; CDR3;
 KW heavy chain variable domain; VH; complementarity determining region 3;
 KW MC-1; antibody.
 XX
 OS Mus sp.
 XX
 PN WO200220615-A2.
 XX
 PD 14-MAR-2002.
 XX
 PF 10-SEP-2001; 2001WO-EP10433.
 XX
 PR 08-SEP-2000; 2000EP-0119694.
 PR 05-SEP-2001; 2001US-0948004.
 XX
 PA (MICR-) MICROMET AG.

KW Alzheimer's; cognitive disorder; schizophrenia; prostate disease;
 KW obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease;
 KW lung disease; thymus disease; digestive disorder; endocrine disorder;
 KW infection; AIDS.
 XX Homo sapiens.
 OS
 FN WO9922243-AL.
 XX
 XX
 PD 06-MAY-1999.
 XX
 XX 23-OCT-1998; 98WO-US22376.
 XX
 XX 24-OCT-1997; 97US-0063387.
 PR 24-OCT-1997; 97US-0062784.
 PR 24-OCT-1997; 97US-0063088.
 PR 24-OCT-1997; 97US-0063089.
 PR 24-OCT-1997; 97US-0063090.
 PR 24-OCT-1997; 97US-0063091.
 PR 24-OCT-1997; 97US-0063092.
 PR 24-OCT-1997; 97US-0063097.
 PR 24-OCT-1997; 97US-0063098.
 PR 24-OCT-1997; 97US-0063099.
 PR 24-OCT-1997; 97US-0063100.
 PR 24-OCT-1997; 97US-0063101.
 PR 24-OCT-1997; 97US-0063109.
 PR 24-OCT-1997; 97US-0063110.
 PR 24-OCT-1997; 97US-0063111.
 PR 24-OCT-1997; 97US-0063148.
 PR 24-OCT-1997; 97US-0063386.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Brewer LA, Carter KC, Duan DR, Ebner R, Endress GA;
 PI Feng P, Florence C, Florence KA, Greene JM, Janat F;
 PI Kaye H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
 PI Ruben SM, Shi Y, Soppet DR, Wei Y, Young P;
 XX
 XX WPI; 1999-303069/25.
 DR
 DR
 DR
 PT New isolated human genes and the secreted polypeptides they encode
 XX
 XX Disclosure; Page 472; 546pp; English.
 XX
 CC The specification describes human secreted proteins. The polynucleotides
 CC and their corresponding secreted polypeptides are useful for preventing,
 CC treating or ameliorating medical conditions, e.g. by protein or gene
 CC therapy. Pathological conditions can also be diagnosed by determining
 CC the amount of the polypeptides in a sample or by determining the presence
 CC of mutations in the polynucleotides. Specific uses are described for each
 CC of the polynucleotides, based on which tissues they are most highly
 CC expressed in, and include developing products for the diagnosis or
 CC treatment of cancer, tumours, neurodegenerative disorders, developmental
 CC abnormalities and fetal deficiencies, blood disorders, leukemias,
 CC diseases of the immune system, autoimmune diseases, hepatic and renal
 CC diseases, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's
 CC and cognitive disorders, schizophrenia, prostate diseases, obesity,
 CC disorders involving osteoclasts such as osteoporosis, arthritis or
 CC malignancies, diseases of testes, lung or thymus, digestive/endocrine
 CC disorders, infections and AIDS. The polypeptides are also useful for
 CC identifying their binding partners.
 XX
 SQ Sequence 22 AA;
 Query Match 33.3%; Score 32; DB 20; Length 22;
 Best Local Similarity 36.4%; Pred. No. 2.2e+02;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 OY 4 EAQICQIIEY 14
 ||| : :
 Db 4 EASVCMHLSY 14

RESULT 9
 AAR92123
 ID AAR92123 standard; peptide; 25 AA.
 XX
 AC AAR92123;
 XX
 XX 15-AUG-1996 (first entry)
 DT
 XX
 XX Streptomyces roseiscleroticus xylanase, xyl4, N-terminal peptide.
 DE
 XX Wood pulp bleaching; xylanase; chromophores; toxic waste reduction;
 KW neutral pH; kraft pulp; paper.
 KW
 XX Streptomyces roseiscleroticus NRRLB-11019.
 OS
 XX US5498534-A.
 PN
 XX 12-MAR-1996.
 PD
 XX
 XX 25-MAR-1992; 92US-0857060.
 PF
 XX 08-JUN-1994; 94US-0257965.
 PR 25-MAR-1992; 92US-0857060.
 PR 30-MAY-1995; 95US-0453289.
 PR
 XX (USDA) US SEC OF AGRIC.
 PA
 XX Elegir G, Grabski AC, Jeffries TW, Patel RN, Szakacs G;
 PI
 XX WPI; 1996-159688/16.
 DR
 XX Method for removing colour from kraft wood pulps - using xylanase
 PT isolated from Streptomyces roseiscleroticus
 PT
 XX Disclosure; Fig 6; 47pp; English.
 PS
 XX AAR92120-R92123 are N-terminal peptides of four xylanases, xyl1, xyl2,
 CC xyl3 and xyl4, encoded by Streptomyces roseiscleroticus strain
 CC NRRLB-11019. The xylanases may be used in a method for bleaching
 CC wood pulp which involves treating wood pulp with xylanase to
 CC release chromophores and extracting it to remove these chromophores.
 CC The wood pulp used is pref. kraft pulp from either soft or hard
 CC wood. Extraction of the chromophores may be performed using an
 CC alkali pref. hydrogen peroxide. Using xylanase as a bleaching
 CC agent reduces the need for using products such as chlorine which
 CC result in the formation of toxic degradation products which are
 CC difficult to remove by conventional waste treatment. Xylanase also
 CC works at a neutral pH, so a large pH shift of the pulp is not
 CC necessary.
 CC
 XX Sequence 25 AA;
 Query Match 33.3%; Score 32; DB 17; Length 25;
 Best Local Similarity 75.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 9 QQIEYIFG 16
 ||| |||
 Db 10 QQSGYIFG 17
 RESULT 10
 AAW80367
 ID AAW80367 standard; peptide; 25 AA.
 XX
 AC AAW80367;
 XX
 XX 13-JAN-1999 (first entry)
 DT
 XX N-terminal sequence of xylanase 4 (xyl4) of S. roseiscleroticus.
 DE
 XX Xylanase; colour removal; wood pulp; kraft pulp; softwood; hardwood.
 KW
 XX

SQ Sequence 14 AA;
 Query Match 33.3%; Score 32; DB 22; Length 14;
 Best Local Similarity 50.0%; Pred. NO. 1.3e+02;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 QQIEYYFGDF 18
 | | : | : |
 DB 2 QDIHFYGRF 11

RESULT 6
 AAW42900
 ID AAW42900 standard; peptide; 20 AA.
 XX AAW42900;
 AC
 XX 28-APR-1998 (first entry)
 DT
 XX Immunogenic Hepatitis A virus peptide YK-1247.
 DE
 XX Immunogenic peptide; immunogenic epitope; VP3 protein;
 KW immune response; antibody.
 KW
 XX Synthetic.
 OS
 XX Hepatitis A virus.
 XX
 XX W09740147-A1.
 PN
 XX 30-OCT-1997.
 PD
 XX
 XX 18-APR-1997; 97WO-US06891.
 PF
 XX 19-APR-1996; 96US-0015644.
 PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Fields HA, Khudyakov YE;
 PI
 XX WPI; 1997-535831/49.
 DR
 XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
 PT immune response to HAV in a mammal or to detect the presence of
 PT antibodies against HAV in a mammal
 PT
 XX Claim 8; Page 110; 140pp; English.
 PS
 XX Peptides AAW42894-905 are immunogenic peptides corresponding to
 CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
 CC substantially similar to a portion of the amino acid sequence of the VP3
 CC protein of HAV corresponding to amino acids 246-491. The present peptide
 CC is derived from amino acids 341-360 and has a reactivity of 19.6% with
 CC acute sera. Compositions containing the peptides can be used to induce an
 CC immune response to HAV in a mammal. The peptides can also be used to
 CC detect the presence of antibodies against HAV in mammalian serum. The
 CC peptides can also be used to make an antibody against HAV by
 CC administering the peptide to a mammal.
 XX

SQ Sequence 20 AA;
 Query Match 33.3%; Score 32; DB 18; Length 20;
 Best Local Similarity 45.2%; Pred. NO. 2e+02;
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 AQICQOIEYFGD 17
 | | | : |
 DB 7 ASICQMFCFWRGD 19

RESULT 7
 AAB69417
 ID AAB69417 standard; Peptide; 21 AA.
 XX

AC AAB69417;
 XX
 XX 20-APR-2001 (first entry)
 DT
 XX Synthetic HAV VP3 peptide, SEQ ID NO: 17.
 DE
 XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KW antigen; major structural capsid polypeptide; HAV antibody detection.
 KW
 XX Hepatitis A virus.
 OS
 XX Synthetic.
 XX
 XX W0200105824-A2.
 PN
 XX 25-JAN-2001.
 PD
 XX 14-JUL-2000; 2000WO-US19267.
 XX
 XX 15-JUL-1999; 99US-0144412.
 PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX Fields HA, Khudyakov YE;
 PI
 XX WPI; 2001-112681/12.
 DR
 XX Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines -
 PT
 XX Claim 7; Page 78; 130pp; English.
 PS
 XX The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum, to
 CC detect the presence of HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC Igm antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy
 CC end of the peptides enhances the Igm antibody reactivity.
 XX

SQ Sequence 21 AA;
 Query Match 33.3%; Score 32; DB 22; Length 21;
 Best Local Similarity 45.2%; Pred. NO. 2.1e+02;
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 AQICQOIEYFGD 17
 | | | : |
 DB 7 ASICQMFCFWRGD 19

RESULT 8
 AAY19621
 ID AAY19621 standard; Protein; 22 AA.
 XX
 XX AAY19621;
 AC
 XX 14-JUL-1999 (first entry)
 DT
 XX
 XX SEQ ID NO 339 from W09922243.
 DE
 XX Human secreted protein; cancer; tumour; neurodegenerative disorder;
 KW developmental abnormality; fetal deficiency; blood disorder; leukemia;
 KW immune system disease; autoimmune disease; hepatic disease;
 KW renal disease; lymphoma; inflammation; allergy; ischemic shock;

CC ligand-binding domain of human monoclonal antibody A6, isolated by
 CC panning against M2 immunoglobulin. Variants of a parental ligand
 CC binding molecule (PLBM) comprise an immunoglobulin (Ig) VH fragment
 CC including the framework (FR) regions of the wild-type/modified Ig VH
 CC domain of human monoclonal antibody A6. The variant differs from PLBM at
 CC residues constituting part of one of the CDRs of PLBM. A library
 CC expressing binding domains can be created by cloning a parental DNA
 CC sequence that encodes a parental domain, in order to produce parental
 CC clones, and replacing a variable region of the clones with a variant DNA
 CC sequence. A library of genetic packages can then be generated, each
 CC package having a surface and a surface protein binding domain expressed
 CC by the variant DNA. Combinatorial libraries are useful for generating
 CC forms of recombinant antibody fragments that specifically bind target
 CC ligands including natural and synthetic molecules and macromolecules to
 CC be used in diagnostic and/or therapeutic processes. Examples include
 CC enzymes, cell-receptors, carbohydrates, lipids, and organic-based and
 CC inorganic-based molecules such as antibiotics, steroids, hormones,
 CC pesticides, herbicides, dyes and polymers.

XX SQ Sequence 23 AA;

Query Match 35.4%; Score 34; DB 22; Length 23;

Best Local Similarity 45.5%; Pred. No. 1.1e+02;

Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 CQQIEYFGDF 18

I : : I I I

Db 11 CFWDVDFKFSDF 21

RESULT 4

AAW93362

ID AAW93362 standard; Protein; 18 AA.

AC AAW93362;

XX 28-MAY-1999 (first entry)

DT Human DPD protein fragment #1.

DE DPD; dihydropyrimidine-dehydrogenase; monoclonal antibody; MAB;

KW immunoassay reagent; cancer patient; treatment; antitumor agent;

KW 5-fluorouracil; affinity purification; toxicity.

XX Homo sapiens.

OS DE19837391-Al.

PN 25-FEB-1999.

PD 18-AUG-1998; 98DE-1037391.

PF 22-AUG-1997; 97EP-0114630.

PR (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Hasegawa M, Yoshikubo T;

PI WPI; 1999-155202/14.

DR Monoclonal antibody specific for dihydropyrimidine dehydrogenase -

PT for assessing patient response to 5-fluorouracil antitumor agents

XX Disclosure; Page 24; 34pp; German.

PS This invention describes a monoclonal antibody (MAB) specific for

XX dihydropyrimidine dehydrogenase (DPD). This MAB is used as immunoassay

CC reagents to identify a lack of DPD in a patient and to assess the

CC sensitivity of cancer patients to treatment with antitumor agents of the

CC 5-fluorouracil (5-FU) type. The MAB can also be used for affinity

CC purification of DPD. DPD is involved in reduction of 5-FU (and related

CC catabolites and derivatives) and lack of it is associated with increased

CC toxicity of this type of antitumor agent. It has specific binding

CC interaction. The MAB provide a sensitive and reliable test for DPD,

CC which is simple, rapid and suitable for routine screening.

XX SQ Sequence 18 AA;

Query Match 34.4%; Score 33; DB 20; Length 18;

Best Local Similarity 50.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 8 CQQIEYFGDF 17

I : : I I I

Db 1 CEKLENNFGD 10

RESULT 5

AAW97227

ID AAW97227 standard; Peptide; 14 AA.

XX AAW97227;

XX 24-JAN-2002 (first entry)

DT Human peptide #502 encoded by a SNP oligonucleotide.

DE Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;

XX neuroprotective; antimicrobial; gene therapy; vaccine; cancer;

KW amyloid protein; angiotensin; apoptosis related protein; cadherin;

KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;

KW complement related protein; cytochrome; kinesin; cytokine; interferon;

KW interleukin; G-protein coupled receptor; thioesterase; inflammation;

KW multifactorial disease; autoimmune disease; infection;

XX nervous system disease.

XX Homo sapiens.

OS WO200147944-A2.

PN 05-JUL-2001.

PD 28-DEC-2000; 2000WO-US35498.

PF 28-DEC-1999; 99US-0173419.

XX 27-DEC-2000; 2000US-0173419.

PR (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

PI WPI; 2001-465210/50.

DR Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,

XX oncogenes and histones, useful for diagnosing and treating, e.g.

PT cancer, autoimmune diseases and infections -

XX Disclosure; Page 3778; 4143pp; English.

PS The present invention relates to oligonucleotides (see AAL26793-AAL34659)

XX encoding polymorphic variants of proteins related to amylases, amyloid

CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,

CC polymerase, oncogenes, histones, kinases, colony stimulating factors,

CC complement related proteins, cytochromes, kinesins, cytokines,

CC interferons, interleukins, G-protein coupled receptors and thioesterases.

CC The present sequence is a peptide encoded by one such oligonucleotide.

CC The oligonucleotides and the peptides encoded by them may be used in the

CC prevention, diagnosis and treatment of diseases associated with

CC inappropriate expression of the proteins listed above. Disorders that may

CC be prevented, diagnosed and/or treated include multifactorial diseases

CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid

CC arthritis, multiple sclerosis, diabetes, systemic lupus erythromatosus

CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,

CC brain, breast, colon and kidney, leukaemia), diseases of the nervous

CC system and an infection of pathogenic organisms.

XX

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:27:28 ; Search time 10.3146 Seconds
(without alignments)
139.834 Million cell updates/sec

Title: US-09-836-073-18

Perfect score: 90
Sequence: 1 VSKLEASTIRQEVYFGDA 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match: 100%
Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
- 2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pap.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
- 7: /cgn2_6/ptodata/2/pubpaa/PTCT_PUBCOMB.pap.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	90	100.0	18	9 US-09-836-073-18	Sequence 18, Appl
2	44.5	49.4	16	9 US-09-836-073-19	Sequence 19, Appl
3	43.5	48.3	18	9 US-09-836-073-2	Sequence 2, Appl
4	43.5	48.3	460	9 US-10-102-806-695	Sequence 695, App
5	43	47.8	184	10 US-09-734-017A-32	Sequence 32, Appl
6	42.5	47.2	17	9 US-09-836-073-13	Sequence 13, Appl
7	42.5	47.2	18	9 US-09-836-073-1	Sequence 1, Appl
8	42.5	47.2	18	9 US-09-836-073-9	Sequence 9, Appl
9	42.5	47.2	18	9 US-09-836-073-14	Sequence 14, Appl
10	42.5	47.2	39	9 US-09-843-676-26	Sequence 26, Appl
11	42.5	47.2	39	9 US-09-766-253-26	Sequence 26, Appl
12	42.5	47.2	39	9 US-09-438-486-26	Sequence 26, Appl
13	42.5	47.2	39	9 US-10-053-758-26	Sequence 26, Appl
14	42.5	47.2	39	9 US-10-054-295-26	Sequence 26, Appl
15	42.5	47.2	39	9 US-10-054-611-26	Sequence 26, Appl
16	42	46.7	19	9 US-09-836-073-16	Sequence 16, Appl
17	42	46.7	37	9 US-09-843-676-24	Sequence 24, Appl
18	42	46.7	37	9 US-09-766-253-24	Sequence 24, Appl
19	42	46.7	37	9 US-09-438-486-24	Sequence 24, Appl

20	42	46.7	37	9 US-10-053-758-24	Sequence 24, Appl
21	42	46.7	37	9 US-10-054-295-24	Sequence 24, Appl
22	42	46.7	37	9 US-10-054-611-24	Sequence 24, Appl
23	40	44.4	876	10 US-09-815-242-5432	Sequence 5432, Ap
24	40	44.4	877	10 US-09-815-242-12556	Sequence 12556, A
25	39	43.3	550	9 US-09-738-626-6950	Sequence 6950, Ap
26	39	43.3	550	9 US-10-226-136-19	Sequence 19, Appl
27	39	43.3	866	12 US-10-007-693-113	Sequence 113, Appl
28	38.5	42.8	18	9 US-09-836-073-4	Sequence 4, Appl
29	38	42.2	58	10 US-09-864-761-33415	Sequence 33415, A
30	38	42.2	212	10 US-09-815-242-10149	Sequence 10149, A
31	38	42.2	487	10 US-09-938-803-6	Sequence 6, Appl
32	38	42.2	883	10 US-09-815-242-13382	Sequence 13382, A
33	38	42.2	883	10 US-09-815-242-13684	Sequence 13684, A
34	37	41.1	18	9 US-09-836-073-5	Sequence 5, Appl
35	37	41.1	18	9 US-09-836-073-15	Sequence 15, Appl
36	37	41.1	38	9 US-09-843-676-25	Sequence 25, Appl
37	37	41.1	38	9 US-09-766-253-25	Sequence 25, Appl
38	37	41.1	38	9 US-09-438-486-25	Sequence 25, Appl
39	37	41.1	38	9 US-10-053-758-25	Sequence 25, Appl
40	37	41.1	38	9 US-10-054-295-25	Sequence 25, Appl
41	37	41.1	38	9 US-10-054-611-25	Sequence 25, Appl
42	37	41.1	88	9 US-09-908-739-24	Sequence 24, Appl
43	37	41.1	88	9 US-09-908-739-47	Sequence 47, Appl
44	37	41.1	260	10 US-09-925-301-881	Sequence 881, Appl
45	37	41.1	268	9 US-09-557-796-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-09-836-073-18
; Sequence 18, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; PRIOR FILING DATE: 2002-10-24
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; TYPE: PRT
; LENGTH: 18
; ORGANISM: Mosquito
US-09-836-073-18

Query Match 100.0%; Score 90; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSKLEASTIRQEVYFGDA 18
Db 1 VSKLEASTIRQEVYFGDA 18

RESULT 2
US-09-836-073-19
; Sequence 19, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila
US-09-836-073-19

Query Match 49.4%; Score 44.5; DB 9; Length 16;
Best Local Similarity 71.4%; Pred. No. 0.23;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 5 EASTIRO-EYYFGD 17
| : ||| |||||
Db 2 ERAIRQVEYFGD 15

RESULT 3

US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Baldiva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-073-2

Query Match 48.3%; Score 43.5; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.39;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 4 LEASTIRO-EYYFGD 17
||| : |||||
Db 3 LEAQCQIEYFGD 17

RESULT 4

US-10-102-806-695
; Sequence 695, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 695
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-102-806-695

Query Match 48.3%; Score 43.5; DB 9; Length 460;
Best Local Similarity 55.6%; Pred. No. 12;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEASTIRO-EYYFGD 17
: : ||| : |||||
Db 62 MAALAKICHQIEYFGD 79

RESULT 5

US-09-734-017A-32
; Sequence 32, Application US/09734017A
; Patent No. US20020142422A1
; GENERAL INFORMATION:
; APPLICANT: Lerchl, Jens
; APPLICANT: Renz, Andreas
; APPLICANT: Ehrhardt, Thomas
; APPLICANT: Reindl, Andreas
; APPLICANT: Cirpus, Petra
; APPLICANT: Bischoff, Friedrich
; APPLICANT: Frank, Markus
; APPLICANT: Freund, Annette
; APPLICANT: Duwenig, Elke
; APPLICANT: Schmidt, Ralf-Michael
; TITLE OF INVENTION: Moss genes from Physcomitrella patens encoding proteins invol
; TITLE OF INVENTION: the
; TITLE OF INVENTION: synthesis of amino acids, vitamins, cofactors, nucleotides a
; TITLE OF INVENTION: nucleosides
; FILE REFERENCE: BASF-NAE-1331-99-US
; CURRENT APPLICATION NUMBER: US/09/734,017A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/171,100
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1/WordPerfect
; SEQ ID NO 32
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-734-017A-32

Query Match 47.8%; Score 43; DB 10; Length 184;
Best Local Similarity 41.2%; Pred. No. 5.7;
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 VSKLEASTIROEYFGD 17
: : ||| : |||||
Db 35 IHSKAQTMLEYYIGD 51

RESULT 6

US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baldiva, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match 47.2%; Score 42.5; DB 9; Length 17;
Best Local Similarity 66.7%; Pred. No. 0.59; Length 18;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 4 LEASTIRQ-EYFQD 17
Db 2 LEAKICHQIEYFQD 16

RESULT 7

US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match 47.2%; Score 42.5; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.59; Length 18;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 4 LEASTIRQ-EYFQD 17
Db 3 LEAKICHQIEYFQD 17

RESULT 8

US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match 47.2%; Score 42.5; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.59; Length 18;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 4 LEASTIRQ-EYFQD 17
Db 3 LEAKICHQIEYFQD 17

RESULT 9

US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14

Query Match 47.2%; Score 42.5; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.59; Length 18;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 4 LEASTIRQ-EYFQD 17
Db 3 LEAKICHQIEYFQD 17

RESULT 10
US-09-843-676-26
; Sequence 26, Application US/09843676
; Patent No. US20020164786A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
; TITLE OF INVENTION: No. US20020164786A1 Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/843,676
; FILING DATE: 26-Apr-2001
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-00293005

Qy 9 IRQ-EYFGDA 18
Db 2 LRQVEYFGDA 12
RESULT 12
US-09-438-486-26

US-09-438-486-26
; Sequence 26, Application US/09438486
; Publication NO. US20030009019A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.

51;

APPLICANT: Andrews, William H.
TITLE OF INVENTION: NO. US20030009019A1el Telomerase
NUMBER OF SEQUENCES: 223
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111-3834

```

;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/438,486
; FILING DATE: 12-NOV-1999
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997

```

/ FILING DATE: 06-MAY-1997
 / CLASSIFICATION: 536
 / PRIOR APPLICATION DATA:
 / PRIOR APPLICATION NUMBER: US 08/846,017
 / FILING DATE: 25-APR-1997
 / CLASSIFICATION: 536
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 08/844,419
 / FILING DATE: 18-APR-1997
 / CLASSIFICATION: 536
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: US 08/724,643
 / FILING DATE: 01-OCT-1996
 / CLASSIFICATION: 536
 / ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002931US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ. ID NO. 26:
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-09-438-486-26

```

Query Match 47.28; Score 42.5; DB 9; Length 39;


```

; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
;
; TITLE OF INVENTION: NO. US20030059787A1el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,611
; FILING DATE: 18-Jan-2002
; CLASSIFICATION: 536
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/854,050
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
; INFORMATION FOR SEQ ID NO: 26:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030059787A1 Relevant
; TOPOLOGY: No. US20030059787A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
;
; US-10-054-611-26

```

```

Query Match 47.2%; Score 42.5; DB 9; Length 39;
Best Local Similarity 81.8%; Pred. NO. 1.3;
Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

```

```

QY 9 IRQ-EYFGDA 18
Db 2 LRQVEYFGDA 12

```

```

Search completed: April 23, 2003, 13:38:23
Job time : 11.3146 secs

```



```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3
    Query Match          47.2%; Score 42.5; DB 4; Length 18;
    Best Local Similarity 56.7%; Pred. No. 0.23;
    Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY      4 LEASTIRQ-EYVFGD 17
       ||| | | | | | |
Db      3 LEAKICHQIEYVFGD 17

RESULT 3
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
; OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4
    Query Match          47.2%; Score 42.5; DB 4; Length 18;
    Best Local Similarity 56.7%; Pred. No. 0.23;
    Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY      4 LEASTIRQ-EYVFGD 17
       ||| | | | | | |
Db      3 LEAKICHQIEYVFGD 17

RESULT 4
US-08-851-843A-26
; Sequence 26, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
```

```
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-851-843A-26
    Query Match          47.2%; Score 42.5; DB 3; Length 39;
    Best Local Similarity 81.8%; Pred. No. 0.58;
    Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY      9 IRQ-EYVFGDA 18
       :| | | | | | | |
Db      2 LKQVEYVFGDA 12

RESULT 5
US-08-974-549A-216
; Sequence 216, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 216:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-549A-216

```

```

Query Match 47.2%; Score 42.5; DB 4; Length 39;
Best Local Similarity 81.8%; Pred. No. 0.58;
Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

```

```

QY 9 IRQ-EYYFGDA 18
Db 2 LRQVEYFGDA 12

```

```

RESULT 6
US-08-854-050-26
; Sequence 26, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin

```

```

; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-854-050-26

```

```

Query Match 47.2%; Score 42.5; DB 4; Length 39;
Best Local Similarity 81.8%; Pred. No. 0.58;
Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

```

```

QY 9 IRQ-EYYFGDA 18
Db 2 LRQVEYFGDA 12

```

```

RESULT 7
US-09-430-323-26
; Sequence 26, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase

```

```
;
;
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-OCT-1999
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-430-323-26
;
;
; Query Match 47.2%; Score 42.5; DB 4; Length 39;
; Best Local Similarity 81.8%; Pred. No. 0.58;
; Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
;
; Qy 9 IRQ-EYVFGDA 18
; Db 2 LRQVEYVFGDA 12
;
; RESULT 8
; US-08-851-843A-24
; Sequence 24, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
```

```
;
;
; COUNTRY: United States of America
; ZIP: 94111
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
;
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
;
; US-08-851-843A-24
;
; Query Match 46.7%; Score 42; DB 3; Length 37;
; Best Local Similarity 100.0%; Pred. No. 0.67;
; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 11 QEYVFGD 17
; Db 4 QEYVFGD 10
;
; RESULT 9
; US-08-854-050-24
; Sequence 24, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0300
; TELEFAX: (415) 576-0200
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-854-050-24

Query Match 46.7%; Score 42; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QEYIFGD 17
Db 4 QEYIFGD 10

RESULT 10
US-09-430-323-24
; Sequence 24, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; LINGNER, Joachim
; NAKAMURA, Toru
; CHAPMAN, Karen B.
; MORIN, Gregg B.
; HARLEY, Calvin
; ANDREWS, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0300
; TELEFAX: (415) 576-0200
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-430-323-24

Query Match 46.7%; Score 42; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QEYIFGD 17
Db 4 QEYIFGD 10

RESULT 11
US-09-457-040B-24
; Sequence 24, Application US/09457040B
; Patent No. 6387841
; GENERAL INFORMATION:
; APPLICANT: Vertex Pharmaceuticals Incorporated
; APPLICANT: Bellon, Steve
; TITLE OF INVENTION: Crystallized P38 Complexes
; FILE REFERENCE: VPI/98-14
; CURRENT APPLICATION NUMBER: US/09/457,040B
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Yeast
; US-09-457-040B-24

Query Match 44.4%; Score 40; DB 4; Length 546;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 3 KLEAST--IRQEYVFG 16
Db 124 KILAGTPGIQEYVFG 139

RESULT 12
US-08-785-071A-2
; Sequence 2, Application US/08785071A
; Patent No. 576750
; GENERAL INFORMATION:

```

APPLICANT: Hodgson, John
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 5776750e1 tRNA Synthetase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/785,071A
FILING DATE: 17-JAN-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9601099.6
FILING DATE: 19-JAN-1996
APPLICATION NUMBER: 9622617.0
FILING DATE: 27-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimml, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31355-6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 876 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-785-071A-2

Query Match 44.4%; Score 40; DB 1; Length 876;
Best Local Similarity 61.5%; Pred. No. 62;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VSKLEASTIRQY 13
: ||:| |||:
Db 1 MKKLKASEIRQKY 13

RESULT 13
US-09-012-872-2
Sequence 2, Application US/09012872
Patent No. 6060294
GENERAL INFORMATION:
APPLICANT: Hodgson, John
APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 6060294e1 tRNA Synthetase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/012,872
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/785,071
FILING DATE: 17-JAN-1997
APPLICATION NUMBER: 9601099.6
FILING DATE: 19-JAN-1996
APPLICATION NUMBER: 9622617.0
FILING DATE: 27-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimml, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P31355-6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 876 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-012-872-2

Query Match 44.4%; Score 40; DB 3; Length 876;
Best Local Similarity 61.5%; Pred. No. 62;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VSKLEASTIRQY 13
: ||:| |||:
Db 1 MKKLKASEIRQKY 13

RESULT 14
US-08-849-480A-5
Sequence 5, Application US/08849480A
Patent No. 5981184
GENERAL INFORMATION:
APPLICANT: MELCHERS, Klaus
TITLE OF INVENTION: SCREENING MODEL
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN, PLLC
STREET: 400 - 7th Street, N. W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/849,480A
FILING DATE: 02-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/04711
FILING DATE: 30-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P4442970.3
FILING DATE: 02-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 19505645.0
FILING DATE: 18-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: AISENBERG, Irwin M.
REGISTRATION NUMBER: 19,007
REFERENCE/DOCKET NUMBER: 8125/P60984U50

```
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-6666
; TELEFAX: 202/393-5350
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; STRAIN: Helicobacter pylori 69A
; INDIVIDUAL ISOLATE: Clinical isolate 69A
; IMMEDIATE SOURCE:
; LIBRARY: Helicobacter pylori 69A - gene library in
; LIBRARY: vector pRH160
; CLONE: pRH514
; US-08-849-480A-5

Query Match 43.9%; Score 39.5; DB 2; Length 506;
Best Local Similarity 52.6%; Pred. No. 40;
Matches 10; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

Qy 3 KLEASTIR--QYYFQDA 18
|| || :|||||
Db 178 KLPEKTIELNDKEYYFSDA 196

RESULT 15
US-08-674-168-29
; Sequence 29, Application US/08674168
; Patent No. 5804414
; GENERAL INFORMATION:
; APPLICANT: MORIYA, Mika
; APPLICANT: MATSUI, Hiroshi
; APPLICANT: YOKOZAKI, Kenzo
; APPLICANT: HIRANO, Seiko
; APPLICANT: HAYAKAWA, Atsushi
; APPLICANT: IZUL, Masako
; APPLICANT: SUGIMOTO, Masakazu
; TITLE OF INVENTION: METHOD OF AMPLIFYING GENE USING
; TITLE OF INVENTION: ARTIFICIAL TRANSPOSON
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 JEFFERSON DAVIS HIGHWAY, SUITE # 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/674,168
; FILING DATE: 01-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-166541
; FILING DATE: 30-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-810-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR

; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-674-168-29

Query Match 43.3%; Score 39; DB 1; Length 550;
Best Local Similarity 56.2%; Pred. No. 55;
Matches 9; Conservative 2; Mismatches 3; Indels 2; Gaps 1;

Qy 4 LEAS--TIRQYYFQD 17
|||| :|||||
Db 154 LEASGAKVTREYFND 169

Search completed: April 23, 2003, 13:36:35
Job time : 12.2247 secs
```

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:23 ; Search time 28.5169 Seconds
(without alignments)
84.109 Million cell updates/sec

Title: US-09-836-073-18

Perfect score: 90

Sequence: 1 VSKLEASTIRQYFVGD 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52.5	58.3	390	22	ABB65316 Drosophila melanog
2	44	48.9	413	21	AAG38952 Arabidopsis thalia
3	44	48.9	512	22	ABB57841 Drosophila melanog
4	43.5	48.3	92	21	AAG01351 Human secreted pro
5	43.5	48.3	136	23	ABP40135 Staphylococcus epi
6	43.5	48.3	408	17	AAW03716 Human autoantigen
7	43.5	48.3	439	22	ABG08417 Novel human diagno
8	43.5	48.3	460	21	RAA58987 Breast and ovarian
9	43.5	48.3	460	23	ABP41511 Human ovarian anti
10	43	47.8	129	22	AAW84747 Human immune/haema

11	43	47.8	448	14	AAW41516 Cystathione beta-s
12	43	47.8	448	14	AAW40920 NHS-5 Protein. Ye
13	43	47.8	507	14	AAW42284 Protein which lowe
14	43	47.8	1252	22	RAE02008 Yeast cyathionin
15	42.5	47.2	18	21	AAW52200 Human la autoantig
16	42	46.7	935	22	ABB62742 Drosophila melanog
17	41	45.6	980	21	AAG51323 Arabidopsis thalia
18	41	45.6	1007	20	AAW19957 B. burgdorferi ant
19	41	45.6	1031	21	AAG51322 Arabidopsis thalia
20	41	45.6	1036	20	AAW19956 B. burgdorferi ant
21	41	45.6	1277	21	AAG51321 Arabidopsis thalia
22	40	44.4	176	22	ABB64852 Drosophila melanog
23	40	44.4	444	22	AAW54425 Propionibacterium
24	40	44.4	876	18	AAW21898 Alanyl-tRNA synthe
25	40	44.4	876	22	AAU33936 Staphylococcus aur
26	40	44.4	877	22	AAU36963 Helicobacter-speci
27	39.5	43.9	506	17	AAW97281 H. pylori GHP0.265
28	39.5	43.9	506	19	AAW98625 Propionibacterium
29	39	43.3	71	22	AAW64740 Human ORFX protein
30	39	43.3	88	23	ABP01014 Novel human secret
31	39	43.3	93	22	AAU29541 Novel human diagno
32	39	43.3	126	22	ABG00128 Novel human diagno
33	39	43.3	290	22	ABG12104 Novel human diagno
34	39	43.3	324	22	ABG13802 Novel human diagno
35	39	43.3	327	22	ABG04346 Brevibacterium lac
36	39	43.3	550	18	AAW23282 Brevibacterium lac
37	39	43.3	550	19	AAW68149 Brevibacterium lac
38	39	43.3	550	19	AAW69551 B. lactofermentum
39	39	43.3	550	19	AAW47398 C. glutamicum prote
40	39	43.3	550	22	AAG93196 Drosophila melanog
41	39	43.3	717	22	ABB71311 C. pneumoniae Ctl1
42	39	43.3	866	22	AAU38917 C. pneumoniae prot
43	39	43.3	872	20	AAW34745 Drosophila melanog
44	39	43.3	1165	22	ABB57845 Drosophila melanog
45	39	43.3	1165	22	ABB66714 Drosophila melanog

ALIGNMENTS

RESULT 1
ABB65316
ID ABB65316 standard; Protein; 390 AA.
XX
AC ABB65316;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 22740.
XX
KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
XX
OS Drosophila melanogaster.
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
(PERKE) PE CORP NY.
Venter JC, Adams M, Li PWD, Myers EW;
WPI; 2001-656860/75.
N-PSDB; ABL09419.
New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX Disclosure: SEQ ID NO 22740; 2lpp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from *Drosophila*. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABLO1840-ABL16175) and the encoded proteins

XX (ABR57737-ABR72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 390 AA;

XX

XX Query Match 58.38; Score 52.5; DB 22; Length 390;

XX Best Local Similarity 66.78; Pred. No. 0.62;

XX Matches 12; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

XX

QY 2 SKLEASTIRO-EYVFGDA 18

DB 48 TKQERAIHQVEYVFGDA 65

XX : | | | | | | | | | |

XX

RESULT 2

AAG38952

ID AAG38952 standard; Protein; 413 AA.

XX

XX AAG38952;

XX

XX 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 48129.

XX

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX

XX Arabidopsis thaliana.

XX

XX EP1033405-A2.

XX

XX 06-SEP-2000.

XX

XX 25-FEB-2000; 2000EP-0301439.

XX

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132407.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 06-MAY-1999; 99US-0132487.

XX 07-MAY-1999; 99US-0132863.

XX 11-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX

Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VSKLEASTIRQY 14
: | | | | | | | | | |
Db 278 ISPLQAAIRQSY 291

RESULT 4

AAAG01351
ID AAG01351 standard; Protein; 92 AA.

XX AC AAG01351;

XX DT 06-OCT-2000 (first entry).

XX DE Human secreted protein, SEQ ID NO: 5432.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping.

XX OS Homo sapiens.

XX PN EP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-0200610.

XX PR 26-FEB-1999; 99US-0122487.

XX PA (GEST) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI; 2000-500381/45.

XX DR N-PSDB; AAC01357.

XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX PS Claim 13; SEQ ID 5432; 71pp + CD-ROM; English.

XX CC The present sequence is a polypeptide encoded by one of a large number
XX CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
XX CC were prepared from total human RNAs or polyA+ RNAs derived from 30
XX CC different tissues. EST sequences usually correspond mainly to the 3'
XX CC untranslated region (UTR) of the mRNA because they are often obtained
XX CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
XX CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
XX CC those cases where longer cDNA sequences have been obtained, the full 5'
XX CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
XX CC ends and can therefore be used to obtain full length cDNAs and genomic
XX CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
XX CC chromosome mapping procedures. They are used to obtain upstream
XX CC regulatory sequences and to design expression and secretion vectors.

XX SQ Sequence 92 AA;

Query Match 48.3%; Score 43.5; DB 21; Length 92;

Best Local Similarity 55.6%; Pred. No. 4.4;

Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEASTIRQ-EYFFGD 17

Db 10 MAALEAKICHQIEYFFGD 27

RESULT 5

ABP40135

ID ABP40135 standard; Protein; 136 AA.

XX AC ABP40135;

XX DT 24-JUL-2002 (first entry)

XX KW Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:4980.

XX KW Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX KW antibacterial; gene therapy.

XX OS Staphylococcus epidermidis.

XX PN US6380370-B1.

XX PD 30-APR-2002.

XX PF 13-AUG-1998; 98US-0134001.

XX PR 14-AUG-1997; 97US-055779P.

XX PR 08-NOV-1997; 97US-064964P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Doucette-Stamm LA, Bush D;

XX DR WPI; 2002-381255/41.

XX DR N-PSDB; ABN92680.

XX PT Novel isolated nucleic acid encoding a Staphylococcus epidermidis
XX PT polypeptide, useful for diagnosing and treating bacterial infections -

XX PS Disclosure; SEQ ID 4980; 267pp; English.

XX CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
XX CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
XX CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
XX CC antibacterial activity and can be used in gene therapy. The sequences
XX CC can also be used in the diagnosis and treatment of bacterial infections,
XX CC particularly S. epidermidis infections. The sequences can be used to
XX CC screen for compounds able to interfere with the S. epidermidis life
XX CC cycle or inhibit S. epidermidis infection.

XX CC N.B. The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from the
XX CC USPTO web site.

XX SQ Sequence 136 AA;

Query Match 48.3%; Score 43.5; DB 23; Length 136;

Best Local Similarity 57.9%; Pred. No. 7.2;

Matches 11; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEAST-IRQIEYFFGDA 18

Db 98 ISKLRSETDERQVYFFDA 116

RESULT 6

AAW03716

ID AAW03716 standard; protein; 408 AA.

XX AC AAW03716;

XX DT 12-MAR-1997 (first entry)

XX DE Human autoantigen La(SS-B).

XX KW Autoimmune disease; La autoantigen; Sjogren's syndrome;
XX KW systemic lupus erythematosus; diagnosis.

XX OS Homo sapiens.

XX PN US5541291-A.

XX PD 30-JUL-1996.

XX XX

PF 31-DEC-1984; 84US-0687908.
 XX
 PR 27-MAY-1987; 87US-0054871.
 PR 31-DEC-1984; 84US-0687908.
 XX
 PA (UYDU-) UNIV DUKE.
 XX
 XX Keene JD;
 XX
 XX WPI; 1996-362015/36.
 XX
 XX Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
 PT overlap syndrome - useful for diagnosis and treatment of autoimmune
 PT diseases
 XX
 XX Disclosure; Columns 15-16; 21pp; English.
 XX
 XX The human lupus antigen (La) is diagnostic for Sjogren's syndrome,
 CC as well as occurring in systemic lupus erythematosus patients. The
 CC La protein is clinically related to the Ro protein that is highly
 CC common among autoimmune patients. La and Ro antigens sometimes
 CC reside on the same cellular ribonucleoprotein particle; most La
 CC patients contain some Ro antibodies and vice versa. La cDNA has
 CC been isolated from a human liver library.
 XX
 XX
 SQ Sequence 408 AA;
 Query Match 48.3%; Score 43.5; DB 17; Length 408;
 Best Local Similarity 55.6%; Pred. No. 28;
 Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
 QY 1 VSKLEASTIRQ-EYFQD 17
 : : ||| |
 Db 10 MALEAKICHQIEYFQD 27
 RESULT 7
 ABG08417
 ID ABG08417 standard; Protein; 439 AA.
 XX
 AC ABG08417;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #8408.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 XX WO200175067-A2.
 XX
 XX 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US08631.
 XX
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 XX
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS72604.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 XX Claim 20; SEQ ID No 38776; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 439 AA;
 Query Match 48.3%; Score 43.5; DB 22; Length 439;
 Best Local Similarity 55.6%; Pred. No. 30;
 Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
 QY 1 VSKLEASTIRQ-EYFQD 17
 : : ||| |
 Db 40 MALEAKICHQIEYFQD 57
 RESULT 8
 AAB58987
 ID AAB58987 standard; Protein; 460 AA.
 XX
 AC AAB58987;
 XX
 DT 27-MAR-2001 (first entry)
 XX
 DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 695.
 XX
 KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW neoplastic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease.
 XX
 OS Homo sapiens.
 XX
 XX WO2000055173-A1.
 XX
 XX 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-US05881.
 XX
 XX 12-MAR-1999; 99US-0124270.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 XX
 XX WPI; 2000-611515/58.
 DR N-PSDB; AAF21890.
 XX
 XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular

disorders and neurological diseases -

Claim 11; Page 1149-11150; 1299pp; English.

Sequences AAF21614 - AAF22031 represent DNA sequences encoding human proteins AAB58711 - AAB59128. The DNA and protein sequences are associated with breast and ovarian cancer. Included in the invention are sequences AAF22032 - AAF22040 and AAB59129 which are used in the isolation and characterisation of the DNA and protein sequences of the invention. The breast and ovarian cancer associated DNA, protein, agonist or antagonist sequences exhibit cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor; antiparasitic and cardiant activity. The antibacterial; antifungal; antiparasitic and cardiant activity. The polynucleotide and protein sequences are used in the diagnosis of cancer, particularly breast and ovarian cancer. The nucleic acid sequences, proteins, agonists and antagonists may also be used in the diagnosis, prevention and treatment of immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; cardiovascular disorders such as myocardial ischaemia; wound healing; neurological diseases such as cerebral anoxia and epilepsy; and infectious diseases.

Sequence 460 AA;

Query Match 48.3%; Score 43.5; DB 21; Length 460;
Best Local Similarity 55.6%; Pred. No. 32;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEASTIRO-EYFFGD 17
:: ||| | |||||
Db 62 MALEAKICHQIEYFFGD 79

RESULT 9

ABP41511 ID ABP41511 standard; Protein: 460 AA.

AC ABP41511;

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HWAF56, SEQ ID NO:2643.

Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
ovarian cancer; breast cancer; tumour; reproductive system disorder;
infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
inflammatory condition; immune disorder; blood disorder;
cardiovascular disorder; respiratory disorder; neurological disorder;
gastrointestinal disorder; urinary system disorder; drug screening;
gene therapy; chromosome mapping; forensic analysis;
antibody preparation; cytostatic; immunomodulatory; neuroprotective;
antiinflammatory; gynaecological; reproductive.

OS Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US18569.

XX 07-JUN-2000; 2000US-209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI, 2002-147878/19.

XX N-PSDB; ABQ54588.

Isolated nucleic acid molecules encoding novel ovarian polypeptides, useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian cancer), immune disorders, cardiovascular disorders and neurological diseases -

Claim 11; SEQ ID NO 2643; 2922pp; English.

The invention relates to 2175 novel human ovarian antigens (ABP41054-ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various ovary and/or breast-related disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis), and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents a human ovarian antigen of the invention.
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 460 AA;

Query Match 48.3%; Score 43.5; DB 23; Length 460;
Best Local Similarity 55.6%; Pred. No. 32;
Matches 10; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 VSKLEASTIRO-EYFFGD 17
:: ||| | |||||
Db 62 MALEAKICHQIEYFFGD 79

RESULT 10

AAM84747 ID AAM84747 standard; Protein: 129 AA.

XX AAM84747;

XX 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen SEQ ID NO:12340.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis.
XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 27-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0211486.
 PR 30-JUN-2000; 2000US-0211535.
 PR 07-JUL-2000; 2000US-02115647.
 PR 07-JUL-2000; 2000US-02116880.
 PR 11-JUL-2000; 2000US-02117487.
 PR 11-JUL-2000; 2000US-02117496.
 PR 14-JUL-2000; 2000US-02118290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226581.
 PR 22-AUG-2000; 2000US-0228668.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 01-NOV-2000; 2000US-0244674.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 (HUMA-) HUMAN GENOME SCI INC.
 PA
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-483426/52.
 N-PSDB; AAK57528.
 XX
 Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 useful for preventing, diagnosing and/or treating cancers and
 metastasis -
 PT
 Claim 11; SEQ ID NO 12340; 3071pp + Sequence Listing; English.
 PS
 XX

CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK62169
 CC represent sequences used in the exemplification of the present invention.
 XX
 XX

SQ Sequence 129 AA;

Query Match 47.8%; Score 43; DB 22; Length 129;
 Best Local Similarity 62.5%; Pred. No. 8.3;
 Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 KLEASTIRQYFVGD 18
 DB 46 KLEATQIHQYTCRDA 61
 ||||| : |||
 ||||| : |||

RESULT 11

AAK41516
 ID AAK41516 standard; Protein; 448 AA.

XX AC AAR41516;

XX DT 03-MAR-1994 (first entry)

DE Cystathione beta-synthase.

XX KW Cystathione beta-synthase; NHS5; hydrogen sulphide; cysteine;
 KW transformation.

XX OS Saccharomyces cerevisiae.

XX PN JP05199863-A.

XX PD 10-AUG-1993.

XX PF 08-MAR-1991; 91JP-0067570.

XX PR 08-MAR-1991; 91JP-0067570.

XX PA (ASAK) ASAKI BREWERIES LTD.

XX DR WPI: 1993-284674/36.

DR N-PSDB; AAK48408.

XX PT Bacteria producing cysteine in high yield - obtd. by
 PT transplanting gene with improved cystathionine beta-synthase
 PT activity into bacteria cell

PS Disclosure; Page 2-5; 12pp; Japanese.

XX CC The protein was prepared from the Saccharomyces cerevisiae hydrogen
 CC sulphide controlling gene (NHS5). Bacteria transformed with this
 CC gene produce cysteine and its derivative glutathione in higher yields.

XX SQ Sequence 448 AA;

Query Match 47.8%; Score 43; DB 14; Length 448;
 Best Local Similarity 58.8%; Pred. No. 38;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SKLEASTIRQYFVGD 18
 DB 292 SKLEASTTKYADVFGNA 308
 ||||| : |||
 ||||| : |||

RESULT 12

AAK40920
 ID AAK40920 standard; Protein; 448 AA.

XX AC AAR40920;

XX DT 18-FEB-1994 (first entry)

XX DE NHS-5 Protein.

XX KW NHS-5; inhibit; yeast; hydrogen sulphide; brewing.

XX OS Yeast.

XX PN JP05192155-A.

XX PD 03-AUG-1993.

XX PF 06-MAR-1991; 91JP-0063710.

XX PR 06-MAR-1991; 91JP-0063710.

XX PA (ASAK) ASAKI BREWERIES LTD.

XX DR WPI: 1993-277470/35.

DR N-PSDB; AAK48580.

XX PT Gene inhibiting formation of hydrogen sulphide from yeast - used
 PT for forming beer in short time

XX PS Claim 1; Page 2; 13pp; Japanese.

XX CC A protein from the sequence (AAQ48580) which inhibits hydrogen
 CC sulphide production in yeast. This allows a faster production of
 CC beer with lower levels of hydrogen sulphide.

XX SQ Sequence 448 AA;

Query Match 47.8%; Score 43; DB 14; Length 448;
 Best Local Similarity 58.8%; Pred. No. 38;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 SKLEASTIRQYFVGD 18

DB 292 SKLEASTTKYADVFGNA 308
 ||||| : |||
 ||||| : |||

RESULT 13

AAK42284
 ID AAK42284 standard; Protein; 507 AA.

XX AC AAR42284;

XX DT 09-MAY-1994 (first entry)

XX DE Protein which lowers amount of H2S formed by yeast.

XX KW yeast; fermentation; beer; hydrogen sulphide;

KW Saccharomyces cerevisiae.

XX OS Saccharomyces cerevisiae.

XX PN JP05244955-A.

XX PD 24-SEP-1993.

XX PF 04-MAR-1992; 92JP-0081429.

```

PR 04-MAR-1992; 92JP-0081429.
PA (ASAK ) ASAHI BREWERIES LTD.
XX
XX WPI; 1993-338922/43.
DR N-PSDB; AAQ50112.
XX
XX Gene inhibiting hydrogen sulphide-formation in yeast - used to
PT transform Saccharomyces genus brewing yeast, giving beer with
PT improved flavour
XX
XX Claim 1; Page 2-5; 17pp; Japanese.
PS
XX
XX The gene encoding the protein can be used in the practical
CC production of beer. Saccharomyces cerevisiae strains transformed
CC with the DNA showed an inhibition of hydrogen sulphide production
CC of between 60 and 80% during beer fermentation. The beer flavour
CC was rather improved by using the transformants.
XX
XX Sequence 507 AA;
SQ
Query Match 47.8%; Score 43; DB 14; Length 507;
Best Local Similarity 58.8%; Pred. No. 44;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 2 SKLEASTIRQYVFGDA 18
Db 351 SKLEASTTKYADVFGNA 367
||||||| : |||:

RESULT 14
AAE02008
ID AAE02008 standard; Protein; 1252 AA.
XX
XX AAE02008;
XX
XX 31-JUL-2001 (first entry)
XX
XX Yeast cystathionine beta-synthase (CBS).
XX
XX Yeast; enzymatic cycling assay; cystathionine beta-synthase; CBS;
KW cystathionine beta-lyase; CBL; cyclical interconversion;
KW homocysteine.
XX
XX Saccharomyces cerevisiae.
XX
XX Key Location/Qualifiers
FH Region 1..238
FT /label= Glutathione_S-transferase
FT /note= "GSR fused to the amino-terminal of CBS"
FT Region 239..745
FT /label= CBS-enzyme
FT /note= "Yeast cystathionine beta-synthase"
FT Misc-difference 459 /note= "Encoded by GAG"
FT Misc-difference 460 /note= "Encoded by GGA"
FT Misc-difference 531 /note= "Encoded by ACC"
FT Region 746..1252 /note= "Repeated CBS region"
FT
XX WO200133187-A2.
XX
XX 10-MAY-2001.
XX
XX 02-NOV-2000; 2000WO-US30518.
XX
XX 02-NOV-1999; 99US-0163126.
PR 10-MAY-2000; 2000US-0203349.
PR 01-NOV-2000; 2000US-0704036.
XX
XX (CATC-) CATCH INC.
PA

```

```

XX Kawasaki G, Webb HK, Owens J, Liedtke R, Forest D, Legaz M;
PI Lawson S;
XX
XX WPI; 2001-335847/35.
DR N-PSDB; AAD06034.
XX
XX Enzyme assays comprising using amounts of e.g. pyruvate and ammonia
PT produced during cyclical interconversion of homocysteine and
PT cystathionine to assess amounts of homocysteine and/or cystathionine in
PT a sample
XX
XX Claim 103; Page 75-79; 80pp; English.
PS
XX
XX The present invention relates to an enzymatic cycling assays using
CC cystathionine beta-synthase (CBS) and cystathionine beta-lyase (CBL) to
CC assess the amount of homocysteine and/or cystathionine in a sample by
CC assaying the amount of pyruvate, ammonia, and/or nicotinamide adenine
CC dinucleotide (NAD+) produced during the cyclical interconversion of
CC homocysteine and cystathionine. The homocysteine reacts with L-serine to
CC form cystathionine in the presence of the enzyme CBS. The CBL catalyse
CC the conversion of cystathionine to homocysteine, pyruvate and ammonia.
CC The enzymatic cycling assay is less expensive and provides a higher
CC sample throughput than the diagnostic assays currently available.
CC The present sequence is yeast CBS having an amino terminal GST
CC fusion protein attached as a result of cloning into the bacterial
CC expression vector pGEX6P-2.
XX
XX Sequence 1252 AA;
SQ
Query Match 47.8%; Score 43; DB 22; Length 1252;
Best Local Similarity 58.8%; Pred. No. 1.3e+02;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 2 SKLEASTIRQYVFGDA 18
Db 1096 SKLEASTTKYADVFGNA 1112
||||||| : |||:

RESULT 15
AA52200
ID AAY52200 standard; peptide; 18 AA.
XX
XX AC AAY52200;
XX
XX 14-MAR-2000 (first entry)
XX
XX Human la autoantigen peptide (LAP).
XX
XX La autoantigen; LAP; Internal ribosome entry site; IRES; translation;
KW viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
KW coronavirus; hepatitis virus; rhadovirus; adenovirus; coxsackie virus;
KW parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
KW vesicular stomatitis virus.
XX
XX Homo sapiens.
XX
XX WO9961613-A2.
XX
XX 02-DEC-1999.
XX
XX 21-MAY-1999; 99WO-US11281.
XX
XX 22-MAY-1998; 98US-0086527.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Das S, Dasgupta A;
XX
XX WPI; 2000-062712/05.
XX
XX New yeast inhibitory peptide useful for inhibiting viral protein
PT

```


PT translation and replication -

XX Claim 5; Page 57; 81pp; English.

XX This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (IRNA see AA245200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rhadovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.

XX Sequence 18 AA;

Query Match 47.2%; Score 42.5; DB 21; Length 18;

Best Local Similarity 66.7%; Pred. No. 0.9;

Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 4 LEASTIRQ-EYFQD 17

Db 3 LEAKICHQIEYFQD 17

Search completed: April 23, 2003, 13:27:18
 Job time : 30.5169 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:34:48 ; Search time 20.0225 Seconds
(without alignments)
185.234 Million cell updates

Title: US-09-836-073-18
Perfect score: 90

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 9297

```
Minimum DB seq length: 0
Maximum DB seq length: 25
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : STPREMBL_21.*
1: sp_arches.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_page.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score		Query Match	Length	DB	ID	Description
	Score						
1	30	33.3	24	13	Q9PRM2	Q9prM2 gallus gall	
2	28	31.1	11	4	Q9UC46	Q9uc46 homo sapien	
3	27	30.0	12	4	Q9UNV5	Q9unv5 homo sapien	
4	26	28.9	20	2	P81949	P81949 bacillus il	
5	26	28.9	23	11	Q9QXR2	Q9qxh2 mus musculus	
6	25	27.8	13	8	Q9THR8	Q9thr8 bryopsis sp	
7	25	27.8	25	8	Q9TGB8	Q9tgb8 alnus crisp	
8	25	27.8	25	8	Q9TGB7	Q9tgb7 alnus glutif	
9	25	27.8	25	8	Q9TGB6	Q9tgb6 alnus marit	
10	25	27.8	25	8	Q9TGB5	Q9tgb5 betula alle	
11	25	27.8	25	8	Q9TGB4	Q9tgb4 betula glan	
12	25	27.8	25	8	Q9TGB3	Q9tgb3 betula papy	
13	25	27.8	25	8	Q9TGB2	Q9tgb2 betula verr	
14	25	27.8	25	8	Q9TGB1	Q9tgb1 betula pube	
15	25	27.8	25	8	Q9TGB0	Q9tgb0 corylus ave	
16	25	27.8	25	8	Q9TGA9	Q9tga9 corylus col	

ALIGNMENTS

RESULT 1

Q9PRM2	PRELIMINARY;	PRT;	24 AA.
ID	Q9PRM2		
AC	Q9PRM2;		
DT	01-MAY-2000	(TREMBLrel. 13, Created)	
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)	
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation)	
DE	17 kDa major immunophilin (Fragment).		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Aves; Neognathae; Galliformes;		
OC	Archosauria; Aves; Neognathae; Galliformes;		
OC	Gallus		
OX	NCBI_taxID=9031;		
RN	[1]		
RP	SEQUENCE.		
RX	MEDLINE=94072550; PubMed=7504525;		
RA	Yem A.W., Reardon I.M., Leone J.W., Heinrichs		
RT	"An active FK506-binding domain of 17,000 daltons		
RT	following limited proteolysis of chicken thyrocyte		
RL	Biochemistry 32:12571-12576(1993).		
DR	HSSP; P27124; IROT.		
DR	InterPro; IPR001179; FKBP_pprase.		
DR	Pfam; PF00254; FKBP; 1.		
SO	SEQUENCE 24 AA; 2555 MW; 2B856573666C33HE		

RESULT 2

Q9UC46
ID Q9UC46 PRELIMINARY; PRT; 11 AA.
AC Q9UC46;
DT 01-MAY-2000 (TRENBLrel. 13, Created)

```

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE Neutrophil inhibitor peptide, NIP-POLYMORPHONUCLEAR neutrophil
DE inhibitor peptide.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE-96326114; PubMed-8703476;
RA Cooper J.A.Jr., Culbreth R.R.;
RT "Characterization of a neutrophil inhibitor peptide harvested from
RT human bronchial lavage; homology to influenza A nucleoprotein.";
RL Am. J. Respir. Cell Mol. Biol. 15:207-215(1996).
SQ SEQUENCE 11 AA; 1262 MW; 951A1C3279C9DB45 CRC64;

Query Match 31.18; Score 28; DB 4; Length 11;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 13 YFFGD 17
Db 5 YFFGD 9

RESULT 3
Q9UNV5 PRELIMINARY; PRT; 12 AA.
AC Q9UNV5;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Inosine monophosphatase 2 (Fragment).
GN IMPA2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97463449; PubMed-9322233;
RA Yoshikawa T., Turner G., Esterling L.E., Sanders A.R.,
RA Detera-Wadleigh S.D.;
RT "A novel human myo-inositol monophosphatase gene, IMP.18p, maps to a
RT susceptibility region for bipolar disorder.";
RL Mol. Psychiatry 2:393-397(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-20284187;
RA Yoshikawa T., Padigaru M., Karkera J.D., Sharma M., Berrettini W.H.,
RA Esterling L.E., Detera-Wadleigh S.D.;
RT "Genomic structure and novel variants of myo-inositol monophosphatase
RT 2.";
RL Mol. Psychiatry 5:165-171(2000).
DR EMBL; AF085630; AAD22139.1; -.
DR EMBL; AF085629; AAD22139.1; JOINED.
FT NON_TER 1
FT NON_TER 12
SQ SEQUENCE 12 AA; 1408 MW; D29A162A15172B06 CRC64;

Query Match 30.0%; Score 27; DB 4; Length 12;
Best Local Similarity 62.5%; Pred. No. 3.4e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 9 IRQEVYFG 16
Db 3 VRQLEFG 10

RESULT 4
P81949 PRELIMINARY; PRT; 20 AA.
ID P81949

```

```

AC P81949;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE D-alpha-glycerophosphatase (EC 3.1.3.-) (D-GPASE) (Fragment).
OS Bacillus licheniformis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1402;
RN [1]
RP SEQUENCE.
RX STRAIN-ATCC 9789;
RX PubMed-9439579;
RA Skraly F.A., Cameron D.C.;
RT "Purification and characterization of a Bacillus licheniformis
RT phosphatase specific for D-alpha-glycerophosphate.";
RL Arch. Biochem. Biophys. 349:27-35(1998).
CC -1- COFACTOR: REQUIRES DIVALENT MAGNESIUM, OR TO A LESSER EXTENT
CC MANGANESE, FOR ACTIVITY.
CC -1- PATHWAY: FINAL STEP IN GLYCEROL FORMATION.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- MISCELLANEOUS: OPTIMUM PH IS 7.1.
CC -1- SIMILARITY: PROBABLY BELONGS TO A FAMILY THAT GROUPS DOG1, DOG2,
CC GPPI AND GPP2.
KW Hydrolase; Magnesium; Manganese; Metal-binding.
FT UNSURE 4
FT NON_TER 20
SQ SEQUENCE 20 AA; 2331 MW; EB2D4838608868CF CRC64;

Query Match 28.9%; Score 26; DB 2; Length 20;
Best Local Similarity 28.6%; Pred. No. 9.3e+02;
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VSKLEASTIRQEY 14
Db 7 VSDFDGTISKODFY 20

RESULT 5
Q9QXH2 PRELIMINARY; PRT; 23 AA.
AC Q9QXH2;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Cyclin D1 (Fragment).
GN CCND1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Eto I.;
RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF12040; AAF23491.1; -.
DR MGD; MGI:88313; Ccnd1.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2748 MW; 5DA4617C73E3341F CRC64;

Query Match 28.9%; Score 26; DB 11; Length 23;
Best Local Similarity 55.6%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 EASTIRQEY 13
Db 9 EVETIRRAY 17

RESULT 6
Q9THR8 PRELIMINARY; PRT; 13 AA.
ID Q9THR8

```

```

Db      4 DYYG 8
Query Match      27.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

RESULT 10
Q9TGB5
ID Q9TGB5 PRELIMINARY; PRT; 25 AA.
AC Q9TGB5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3
OS Betula alleghaniensis.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=21017;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080078; AAD50065.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match      27.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
Db      4 DYYG 8
Query Match      27.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

RESULT 11
Q9TGB4
ID Q9TGB4 PRELIMINARY; PRT; 25 AA.
AC Q9TGB4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3
OS Betula glandulosa.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=21018;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080079; AAD50066.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match      27.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
Db      4 DYYG 8
Query Match      27.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

RESULT 12
Q9TGB3
ID Q9TGB3 PRELIMINARY; PRT; 25 AA.
AC Q9TGB3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3
OS Betula papyrifera (Paper birch).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=3507;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080080; AAD50067.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match      27.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
Db      4 DYYG 8
Query Match      27.8%; Score 25; DB 8; Length 25;
Best Local Similarity 60.0%; Pred. No. 1.8e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

RESULT 13
Q9TGB2
ID Q9TGB2 PRELIMINARY; PRT; 25 AA.
AC Q9TGB2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3
OS Betula verrucosa (White birch) (Betula pendula).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=3505;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080081; AAD50068.1;
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.

```

<hr/>					
Q9TGB7	PRELIMINARY;	PRT;	25 AA.		
ID	Q9TGB7				
AC	Q9TGB7				
DT	01-MAY-2000 (TrEMBLrel. 13, Created)				
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)				
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)				
DE	Ribosomal protein small 3 (Fragment).				
GN	RPS3.				
OS	Alnus glutinosa (Alder).				
OG	Mitochondrion.				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
OC	eurosid1 I; Fagales; Betulaceae; Alnus.				
OX	NCBI_TaxID=3517;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99261656; PubMed=10331271;				
RA	Laroche J., Bousquet J.;				
RT	"Evolution of the mitochondrial rps3 intron in perennial and annual				
RT	rt angiosperms and homology to nad5 intron 1.";				
RL	Mol. Biol. Evol. 16:441-452(1999).				
DR	EWEL; AF080076; AAD50063.1; -				
DR	InterPro; IPR001351; Ribosomal_S3.				
DR	Pfam; PF00417; Ribosomal_S3_N; 1.				
KW	Mitochondrion.				
FT	NON_TER	1			
FT	NON_TER	25			
FT	NON_TER	25			
SEQ	SEQUENCE	25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;			
Query Match 27.8%; Score 25; DB 8; Length 25;					
Best Local Similarity 60.0%; Pred. No. 1.8e+03;					
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;					
QY	12 EYFG 16				
DZ	:::::				
DB	4 DYYG 8				
<hr/>					
RESULT 9					
Q9TGB6	PRELIMINARY;	PRT;	25 AA.		
ID	Q9TGB6				
AC	Q9TGB6;				
DT	01-MAY-2000 (TrEMBLrel. 13, Created)				
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)				
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)				
DE	Ribosomal protein small 3 (Fragment).				
GN	RPS3.				
OS	Alnus maritima.				
OG	Mitochondrion.				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
OC	eurosid1 I; Fagales; Betulaceae; Alnus.				
OX	NCBI_TaxID=21015;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99261656; PubMed=10331271;				
RA	Laroche J., Bousquet J.;				
RT	"Evolution of the mitochondrial rps3 intron in perennial and annual				
RT	rt angiosperms and homology to nad5 intron 1.";				
RL	Mol. Biol. Evol. 16:441-452(1999).				
DR	EWEL; AF080077; AAD50064.1; -				
DR	InterPro; IPR001351; Ribosomal_S3.				
DR	Pfam; PF00417; Ribosomal_S3_N; 1.				
KW	Mitochondrion.				
FT	NON_TER	1			
FT	NON_TER	25			
FT	NON_TER	25			
SEQ	SEQUENCE	25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;			
Query Match 27.8%; Score 25; DB 8; Length 25;					
Best Local Similarity 60.0%; Pred. No. 1.8e+03;					
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;					
QY	12 EYFG 16				

KW Mitochondrion. 1 1
 FT NON_TER 25 25
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 3CAA2B3D962A52D8 CRC64;
 Query Match 27.8%; Score 25; DB 8; Length 25;
 Best Local Similarity 60.0%; Pred. No. 1.8e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 12 EYFG 16
 Db 4 DYYG 8
 RESULT 14
 Q9TGB1 PRELIMINARY; PRT; 25 AA.
 AC Q9TGB1
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Ribosomal protein small 3 (Fragment).
 GN RPS3.
 OS Betula pubescens (downy birch).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fagales; Betulaceae; Betula.
 OX NCBI_TaxID=3878;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99261656; PubMed=10331271;
 RA Laroche J., Bousquet J.;
 RT "Evolution of the mitochondrial rps3 intron in perennial and annual
 angiosperms and homology to nad5 intron 1.";
 RL Mol. Biol. Evol. 16:441-452(1999).
 DR EMBL; AF080082; AAD50069.1; -;
 DR InterPro; IPR001351; Ribosomal_S3.
 DR Pfam; PF00417; Ribosomal_S3_N; 1.
 KW Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 3CAA2B3D962A52D8 CRC64;

Query Match 27.8%; Score 25; DB 8; Length 25;
 Best Local Similarity 60.0%; Pred. No. 1.8e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Qy 12 EYFG 16
 Db 4 DYYG 8

RESULT 15
 Q9TGB0 PRELIMINARY; PRT; 25 AA.
 AC Q9TGB0
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Ribosomal protein small 3 (Fragment).
 GN RPS3.
 OS Corylus avellana (European hazel).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fagales; Betulaceae; Corylus.
 OX NCBI_TaxID=13451;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99261656; PubMed=10331271;
 RA Laroche J., Bousquet J.;
 RT "Evolution of the mitochondrial rps3 intron in perennial and annual

RT angiosperms and homology to nad5 intron 1.";
 RL Mol. Biol. Evol. 16:441-452(1999).
 DR EMBL; AF080084; AAD50071.1; -;
 DR InterPro; IPR001351; Ribosomal_S3.
 DR Pfam; PF00417; Ribosomal_S3_N; 1.
 KW Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 3CAA2B3D962A52D8 CRC64;
 Query Match 27.8%; Score 25; DB 8; Length 25;
 Best Local Similarity 60.0%; Pred. No. 1.8e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 12 EYFG 16
 Db 4 DYYG 8

Search completed: April 23, 2003, 13:47:18
 Job time : 20.0225 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:33:13 ; Search time 4.65169 Seconds
(without alignments)
160.495 Million cell updates/sec

Title: US-09-836-073-18

Perfect score: 90

Sequence: 1 VSKLEASTIRQEYFQDA 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1520

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	30.0	23	1 NU05_SOLTU	P80262 solanum tub
2	26	28.9	25	1 NEUU_CANFA	P34962 canis famli
3	25	27.8	21	1 NLA_ACTSP	P33036 acinetobact
4	24	26.7	22	1 AOFI_MOUSE	Q64133 mus musculu
5	22	24.4	15	1 UREL_MORMO	P17337 moriganelia
6	22	24.4	20	1 COXN_THUOB	P80980 thunnus obe
7	22	24.4	23	1 PRO3_DACGL	P18690 dactylis gl
8	21	23.3	9	1 ALLO_CARMA	P81813 carcinus ma
9	21	23.3	14	1 LPF2_ECOLI	P05985 escherichia
10	21	23.3	15	1 YAA3_RHOA	Q02006 rhodopseudo
11	21	23.3	18	1 SPAH_HELAN	P81098 helianthus
12	21	23.3	19	1 HBB2_UROHA	P18992 uromastix h
13	21	23.3	20	1 VR90_BORPE	P81549 bordetella
14	21	23.3	22	1 23KD_BACST	P80166 bacillus st
15	21	23.3	22	1 YHV4_LACHE	P22296 lactobacill
16	21	23.3	24	1 ATPE_MICLU	P80286 micrococcu
17	20	22.2	11	1 TKNA_CHICK	P19850 gallus gall
18	20	22.2	15	1 MCA2_RHOOP	P56870 rhodococcus
19	20	22.2	19	1 OXLA_OPHHA	P81383 ophiophagus
20	19	21.1	8	1 ALI7_CARMA	P81820 carcinus ma
21	19	21.1	9	1 ALI1_CARMA	P81814 carcinus ma
22	19	21.1	10	1 TKU1_UREUN	P40751 urechis uni
23	19	21.1	13	1 BP37_LEUMA	P81754 leucophaea
24	19	21.1	16	1 MLB_SOUAC	P01207 squalus aca
25	19	21.1	16	1 OPBE_SOLTU	P81419 solanum tub
26	19	21.1	17	1 APFI_PAVLU	P28529 pavlova lut
27	19	21.1	19	1 FIBB_VULVU	P14482 vulpes vulp
28	19	21.1	20	1 COG4_CHIOP	P34156 chionocete
29	19	21.1	20	1 TL19_SPIOL	P82799 spinacia ol
30	19	21.1	21	1 TL19_ARATH	P82658 arabidopsis
31	19	21.1	22	1 MYH7_CANFA	P49824 canis famli
32	19	21.1	24	1 CH60_ACICA	P81874 acinetobact
33	19	21.1	25	1 COX6_NEUCR	Q01359 neurospora

```

34 19 21.1 25 1 DNAB_MYCCA P71500 mycoplasma
35 18 20.0 4 1 OCPI_OCTMI P58648 octopus min
36 18 20.0 7 1 ALL2_CARMA P81805 carcinus ma
37 18 20.0 7 1 ALL3_CARMA P81806 carcinus ma
38 18 20.0 9 1 FIBB_THEGE P19342 theropithec
39 18 20.0 11 1 TKNA_ONCMY P28499 oncorhynchu
40 18 20.0 15 1 PGTS_PELAC P80564 pelobacter
41 18 20.0 15 1 UC27_MAIZE P81528 mycobacteri
42 18 20.0 17 1 Alys_MYCPH P35985 canine aden
43 18 20.0 18 1 HEX_ADECU P80627 zea mays (m
44 18 20.0 18 1 UC21_MAIZE P36502 streptococc
45 18 20.0 19 1 DURB_STRGW

```

ALIGNMENTS

```

RESULT 1
NU05_SOLTU STANDARD; PRT; 23 AA.
ID P80262;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase 29 kDa subunit (EC 1.6.5.3)
DE (EC 1.6.99.3) (Complex I-29KD) (CI-29KD) (Fragment).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. Bintje; TISSUE=Tuber;
RX MEDLINE=94124587; Pubmed=8294484;
RA Herz U., Schroeder W., Liddell A., Leaver C.J., Brennicke A.,
RA Grohmann L.;
RT "Purification of the NADH:ubiquinone oxidoreductase (complex I) of
RT the respiratory chain from the inner mitochondrial membrane of
RT Solanum tuberosum."
RL J. Biol. Chem. 269:2263-2269(1994).
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
CC -!- SUBCELLULAR LOCATION: MATRIX SIDE OF THE MITOCHONDRIAL INNER
CC MEMBRANE.
CC PIR; C49732; C49732.
DR Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
KW NON_TER 23
FT SEQUENCE 23 AA; 2653 MW; 4B41B8FF83412F58 CRC64;
SQ

```

Query Match 30.0%; Score 27; DB 1; Length 23;
Best Local Similarity 25.0%; Pred. No. 1.4e+02;
Matches 2; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```

QY 9 IRQEIYFG 16
    ::::|:::|
DB 15 VKRDYFYG 22

```

```

RESULT 2
NEUU_CANFA STANDARD; PRT; 25 AA.
ID NEUU_CANFA
AC P34962;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neuromedin U-25 (NMU-25) [Contains: Neuromedin U-8 (NMU-8)].
GN NMU.
OS Canis familiaris (Dog).

```

```

|| |:::| ||
5 SKFLAATVQAEVPVLD 21

RESULT 4
ID AOA_MOUSE STANDARD; PRT; 22 AA.
Q64133;
AC
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Amino oxidase [flavin-containing] A [EC 1.4.3.4] (Monoamine oxidase)
DE (MAO-A) (Fragment).
GN MAO.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95312871; PubMed=7792602;
RR Cases O., Seif I., Grimby J., Gaspar P., Chen K., Pourain S.,
RA Mueller U., Aguet M., Babinet C., Shih J.C., de Maeyer E.;
RT "Aggressive behavior and altered amounts of brain serotonin and
RT norepinephrine in mice lacking MAO-A.";
RL Science 268:1763-1766(1995).
CC -1- FUNCTION: Catalyzes the oxidative deamination of biogenic and
CC xenobiotic amines and has important functions in the metabolism of
CC neuroactive and vasoactive amines in the central nervous system
CC and peripheral tissues. MAO-A preferentially oxidizes biogenic
CC amines such as 5-hydroxytryptamine (5-HT), norepinephrine and
CC epinephrine (By similarity).
CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
CC H(2)O(2).
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: MONOMER. HOMO- OR HETERODIMER (CONTAINING TWO SUBUNITS OF
CC SIMILAR SIZE). EACH SUBUNIT CONTAINS A COVALENTLY BOUND FLAVIN.
CC ENZYMICALLY ACTIVE AS MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Mitochondrial outer membrane.
CC -1- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S78615; AAB34677.1;
CC DR EMBL; S78606; AAB34677.1; JOINED.
CC DR MGD; MGI:96915; Mboa.
CC KW Oxidoreductase; Flavoprotein; FAD; Transmembrane; Mitochondrion;
CC Catecholamine metabolism; Neurotransmitter degradation.
CC FT NON_TER 1 1
CC FT NP_BIND <1 >22 FAD (ADP PART) (POTENTIAL).
CC FT NON_TER 22
CC FT NON_TER 22
CC SQ SEQUENCE 22 AA; 2188 MW; 333FDFE8F98B8CDE CRC64;

Query Match 26.7%; Score 24; DB 1; Length 22;
Best Local Similarity 38.5%; Pred.No. 4.7e+02;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 VSKLEASTIRQEQ 13
DB 5 ISGLAARKILSEY 17
|| |:::| ||

RESULT 5
UREL_MORMO STANDARD; PRT; 15 AA.
ID UREL_MORMO
AC P17337;
DT 01-AUG-1990 (Rel. 15, Created)

```


DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Urease alpha subunit (EC 3.5.1.5) (Urea amidohydrolase) (Urease 63 kDa
 DE subunit) (Fragment).
 GN UREC.
 OS Morganella morganii (Proteus morganii).
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Morganella.
 OX NCBI_TaxID=582;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=90264298; PubMed=2345135;
 RA Hu L.-T., Nicholson E.B., Jones B.D., Lynch M.J., Mobley H.L.T.;
 RT "Morganella morganii urease: purification, characterization, and
 RT isolation of gene sequences.";
 RL J. Bacteriol. 172:3073-3080(1990).
 CC -1- CATALYTIC ACTIVITY: Urea + H(2)O = CO(2) + 2 NH(3).
 CC -1- COFACTOR: Binds 2 nickel ions per subunit (Potential).
 CC -1- SUBUNIT: (ALPHA, BETA, GAMMA)(3) (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE UREASE FAMILY.
 DR PIR; A35389; A35389.
 DR InterPro; IPR001924; UreaseA.
 DR PROSITE; PS00145; UREASE_2; PARTIAL.
 DR PROSITE; PS01120; UREASE_1; PARTIAL.
 KW Hydrolase; Metal-binding; Nickel.
 FT NON_TER 15 15
 SQ SEQUENCE 15 AA; 1650 MW; 09E27AA54241687B CRC64;
 Query Match 24.4%; Score 22; DB 1; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.9e+02; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 0;
 QY 10 RQY 13
 Db 5 RQY 8
 RESULT 6
 ID COXN_THUOB STANDARD; PRT; 20 AA.
 AC P80980;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Cytochrome c oxidase polypeptide VIIb-heart (EC 1.9.3.1) (Fragment).
 OS Thunnus obesus (Bigeye tuna).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Scombroidei;
 OC Scombridae; Thunnus.
 OX NCBI_TaxID=8241;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Heart;
 RX MEDLINE=97454291; PubMed=9310366;
 RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
 RA Kadenbach B.;
 RT "The subunit structure of cytochrome-c oxidase from tuna heart and
 RT liver.";
 RL Eur. J. Biochem. 248:99-103(1997).
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -1- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 KW Oxidoreductase; Mitochondrion.
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2303 MW; 0A33BD34006E5AA6 CRC64;
 Query Match 24.4%; Score 22; DB 1; Length 20;
 Best Local Similarity 28.6%; Pred. No. 9.4e+02; Indels 2; Gaps 1;
 Matches 4; Conservative 6; Mismatches 2;
 QY 10 RQY 13
 Db 5 RQY 8

QY 6 ASTROCY--YFGD 17
 Db 1 SNTSHQDHLFYGD 14
 RESULT 7
 ID PRO3_DACGL STANDARD; PRT; 23 AA.
 AC F18690;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Prolamin alpha-3 (Fragment).
 OS Dactylis glomerata (Orchard grass) (Cocksfoot grass).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
 OC Poaceae; Dactylis.
 OX NCBI_TaxID=4509;
 RN [1]
 RP SEQUENCE.
 RA Vvedenskaya I.O., Shlyapnikov S.V., Konarev A.V.;
 RT "Characterization of the N-terminal amino acid sequence of alpha-
 RT prolamine from Dactylis glomerata L.";
 RL Biochimica 51:1519-1522(1986).
 DR PIR; S02201; S02201.
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2938 MW; E756CA1217592F39 CRC64;
 Query Match 24.4%; Score 22; DB 1; Length 23;
 Best Local Similarity 60.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;
 Matches 3; Conservative 2; Mismatches 0;
 QY 10 RQY 14
 Db 11 RQY 15
 RESULT 8
 ID AL10_CARMA STANDARD; PRT; 9 AA.
 AC P81813;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 10.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
 OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
 OC Brachyura; Eubrachyura; Portunoidae; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 9 9
 SQ SEQUENCE 9 AA; 963 MW; 372D79CDBA776C7 CRC64;
 Query Match 23.3%; Score 21; DB 1; Length 9;
 Best Local Similarity 66.7%; Pred. No. 1.1e+05; Indels 0; Gaps 0;
 Matches 4; Conservative 0; Mismatches 2;
 QY 11 RQY 16
 Db 3 RQY 8

```

RESULT 9
LPP2_ECOLI
ID LPP2_ECOLI STANDARD; PRT; 14 AA.
AC P06985;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phenylalanyl-tRNA synthetase operon leader peptide (PheST attenuator
DE peptide).
GN PHEM OR PHL OR B1715 OR Z2744 OR ECS2422 OR STY1774.
OS Escherichia coli,
OS Escherichia coli O157:H7, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334, 601;
RN [1]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=85210878; PubMed=3158742;
RA Springer M., Mayaux J.-F., Fayat G., Plumbridge J.A., Graffe M.,
RA Blanquet S., Grunberg-Manago M.;
RT "Attenuation control of the Escherichia coli phenylalanyl-tRNA
RT synthetase operon.";
RL J. Mol. Biol. 181:467-478(1985).
RN [2]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=84090239; PubMed=6317865;
RA Fayat G., Mayaux J.-F., Sacerdot C., Fromant M., Springer M.,
RA Grunberg-Manago M., Blanquet S.;
RT "Escherichia coli phenylalanyl-tRNA synthetase operon region.
RT Evidence for an attenuation mechanism. Identification of the gene for
RT the ribosomal protein L20.";
RL J. Mol. Biol. 171:239-261(1983).
RN [3]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=88163794; PubMed=3126825;
RA Springer M., Graffe M., Mayaux J.-F., Dardel F., Fayat G.,
RA Blanquet S., Grunberg-Manago M.;
RT "Open reading frames in the control regions of the phenylalanyl-tRNA
RT synthetase operon of E. coli.";
RL Biochimie 69:1065-1070(1987).
RN [4]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [5]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller H.L.,
RA Grobeck E.J., Davis N.W., Lim A., Dmalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [6]
RN SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,

```

```

RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [7]
RN SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks R.M., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies K.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M10423; AAA23961.1; -
DR EMBL; V00291; CAA23563.1; -
DR EMBL; M13251; AAA24333.1; -
DR EMBL; AE000266; AAC74785.1; -
DR EMBL; AE005394; AAG56702.1; -
DR EMBL; AF002558; BAB35845.1; -
DR EMBL; M1627271; CAD02016.1; -
DR PIR; S11551; LFECEFS.
DR EcoGene; EG11272; pheM.
KW Leader peptide; Complete proteome.
SQ SEQUENCE 14 AA; 1762 MW; 7D31C48E0060F0D4 CRC64;
Query Match 23.3%; Score 21; DB 1; Length 14;
Best Local Similarity 33.3%; Pred. No. 9.5e+02;
Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 4 LEASTIRQYYF 15
Db 1 MNAAIPRFFYF 12
: | : | : |
: | : | : |
RESULT 10
YAA3_RHOA
ID YAA3_RHOA STANDARD; PRT; 15 AA.
AC Q02006;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in AADR 3' region (Fragment).
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CGA009;
RX MEDLINE=92394882; PubMed=1522059;
RA Dispensa M., Thomas C.T., Kim M.K., Perrotta J.A., Gibson J.,
RA Harwood C.S.;
RT "Anaerobic growth of Rhodopseudomonas palustris on 4-hydroxybenzoate
RT is dependent on AadR, a member of the cyclic AMP receptor protein
RT family of transcriptional regulators.";
RL J. Bacteriol. 174:5803-5813(1992).
CC -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M92426; AAA26091.1; -
 DR PIR; C43334; C43334.
 KW Hypothetical protein.
 FT NON_TER 15 15
 SQ SEQUENCE 15 AA; 1926 MW; 149B01A548D7C202 CRC64;

Query Match 23.3%; Score 21; DB 1; Length 15;
 Best Local Similarity 42.9%; Pred. No. 1e+03;
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 LRQEVYF 15
 :|||
 DB 9 LRARHYF 15

RESULT 11

ID SFAH_HELAN STANDARD; PRT; 18 AA.
 AC P81098;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Seed fatty acyl-ester hydrolase (EC 3.1.1.1) (Fragment).
 OS Helianthus annuus (Common sunflower).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 CC Heliantheae; Helianthus.
 OK NCBI_TaxID=4232;
 RN [1]

RP SEQUENCE.
 RC STRAIN=CV. RUSTICA / VAR. EUROFLOR; TISSUE=Seed;
 RA Belsson F., Gardies A.-M., Telsere M., Ferte N., Noat G.;
 RT "An esterase neosynthesized in post-germinated sunflower seeds is
 related to a new family of lipolytic enzymes."
 RL Plant Physiol. Biochem. 35:761-765(1997).
 RN [2]

RP CHARACTERIZATION.
 RX MEDLINE=95210327; PubMed=7696323;
 RA Telsere M., Borel M., Callot B., Nari J., Gardies A.-M., Noat G.;
 RT "Purification and characterization of a fatty acyl-ester hydrolase
 from post-germinated sunflower seeds."
 RL Biochim. Biophys. Acta 1255:105-112(1995).
 CC -1- FUNCTION: IMPLICATED IN THE BREAKDOWN OF OIL BODY-STORED LIPIDS
 CC DURING POST-GERMINATION.
 CC -1- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O -> an alcohol + a
 CC carboxylic anion.
 CC -1- TISSUE SPECIFICITY: SEED.
 CC -1- DEVELOPMENTAL STAGE: POST-GERMINATION.
 CC -1- SIMILARITY: BELONGS TO THE "GDSL" FAMILY OF LIPOLYTIC ENZYMES.

DR InterPro: IPR001087; Lipase_GDSL.
 DR PROSITE: PS01098; LIPASE_GDSL_SER; PARTIAL.
 KW Hydrolyase; Lipid degradation; Glycoprotein.
 FT AC_SITE 13 13
 FT NON_TER 18 18
 SQ SEQUENCE 18 AA; 2057 MW; 3E9A80EB6548E862 CRC64;

Query Match 23.3%; Score 21; DB 1; Length 18;
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 13 YVFGDA 18
 :|||
 DB 8 FIFGDS 13

RESULT 12

ID HBB2_UROHA STANDARD; PRT; 19 AA.
 AC P18992;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hemoglobin beta-2 chain (Fragment).
 OS Uromastix hardwickii (Indian spiny-tailed lizard).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Uromastycinae;
 CC Uromastix.
 OK NCBI_TaxID=40250;
 RN [1]

RP SEQUENCE.
 RX MEDLINE=84029159; PubMed=6628672;
 RA Nagvi S., Zaidi Z.H., von Bahr-Lindstroem H., Carlquist M.,
 RA Joernvall H.;
 RT "Characterization of hemoglobin from the lizard Uromastix
 hardwickii."
 RL FEBS Lett. 162:290-295(1983).
 CC -1- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
 CC VARIOUS PERIPHERAL TISSUES.
 CC -1- SUBUNIT: HETEROPTETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
 CC -1- TISSUE SPECIFICITY: RED BLOOD CELLS.
 CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
 DR PIR; A05305; A05305.
 DR InterPro: IPR000971; Globin.
 DR PROSITE: PS01033; GLOBIN; PARTIAL.
 KW Heme; Oxygen transport; Transport; Erythrocyte.

FT NON_TER 1 1
 FT NON_TER 19 19
 SQ SEQUENCE 19 AA; 1914 MW; C40AD8EA30019057 CRC64;

Query Match 23.3%; Score 21; DB 1; Length 19;
 Best Local Similarity 75.0%; Pred. No. 1.3e+03;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 YFGD 17
 :|||
 DB 1 FFGD 4

RESULT 13

ID VR90_BORPE STANDARD; PRT; 20 AA.
 AC P81549;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Virulence-associated outer membrane protein VR90 (Fragment).
 GN VR90.

OS Bordetella pertussis.
 CC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 CC Bordetella.
 OK NCBI_TaxID=520;
 RN [1]

RP SEQUENCE.
 RC STRAIN=Tohama I;
 RX MEDLINE=99179239; PubMed=10079522;
 RA Passerini de Rossi B.N., Friedman L.E., Gonzalez Flecha F.L.,
 RA Castello P.R., Franco M.A., Rossi J.P.F.C.;
 RT "Identification of Bordetella pertussis virulence-associated outer
 membrane proteins."
 RL FEBS Microbiol. Lett. 172:9-13(1999).
 KW Outer membrane; Virulence.
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2116 MW; D24E1CDCA665206C CRC64;

Query Match 23.3%; Score 21; DB 1; Length 20;
 Best Local Similarity 36.4%; Pred. No. 1.4e+03;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 KLEASTIQEY 13
:|:|:|
Db 6 QLPSTVEGEY 16

Db 10 DGNSTGQKY 19

Search completed: April 23, 2003, 13:43:53
Job time : 5.75169 secs

RESULT 14

23KD_BACST
ID 23KD_BACST STANDARD; PRT; 22 AA.
AC P80166;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE 23 kDa basic protein (Fragment).
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE.
RA Vorlias C.E.;
RL Submitted (OCT-1992) to the SWISS-PROT data bank.
FT NON_TER 22
SQ SEQUENCE 22 AA; 2376 MW; 96C604E42CE0BF6C CRC64;

Query Match 23.3%; Score 21; DB 1; Length 22;
Best Local Similarity 44.4%; Pred. No. 1.6e+03;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 VSKLEASTI 9
|:|:|:|
Db 9 VSKVDLSEV 17

RESULT 15

YHV4_LACHE
ID YHV4_LACHE STANDARD; PRT; 22 AA.
AC P22296;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in HIV 3' region (ORF4) (Fragment).
OS Lactobacillus helveticus.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus
OX NCBI_TaxID=1587;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=481;
RX MEDLINE=91035244; PubMed=2228964;
RA Joergers M.C., Klaenhammer T.R.;
RT "Cloning, expression, and nucleotide sequence of the Lactobacillus
helveticus 481 gene encoding the bacteriocin helveticin J.";
RL J. Bacteriol. 172:6339-6347(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M59360; AAA63275.1;
DR PIR; D37145; D37145.
KW Hypothetical protein.
FT NON_TER 22
SQ SEQUENCE 22 AA; 2484 MW; 0FA71BE0B086E4F2 CRC64;

Query Match 23.3%; Score 21; DB 1; Length 22;
Best Local Similarity 30.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 EASTIQEY 14

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:36:43 ; Search time 9.40449 Seconds
(without alignments)
183.999 Million cell updates/sec

Title: US-09-836-073-18

Perfect score: 90

Sequence: 1 VSKLEASTIRQYYFGDA 18

Scoring table:

BLOSUM62

Gapop. 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 4984

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	37	41.1	25	2 A49038	Ig lambda chain V
2	35	38.9	25	2 C57001	endo-1,4-beta-xyla
3	34	37.8	12	2 PH0746	T-cell receptor be
4	33	36.7	12	2 PH1466	T-cell receptor be
5	32	35.6	12	2 PH0771	T-cell receptor be
6	31	34.4	12	2 S26559	major immunophilin
7	30	33.3	24	2 B49480	Ig heavy chain V r
8	30	33.3	25	2 PH1716	Ig heavy chain V r
9	28	31.1	12	2 S26549	T-cell receptor be
10	28	31.1	12	2 PH1454	T-cell receptor al
11	28	31.1	18	2 PH1368	Ig heavy chain DJ
12	28	31.1	19	2 PH1629	Ig H chain V-D-J r
13	28	31.1	18	2 S04169	IGA-binding protein
14	28	31.1	25	2 PH1734	Ig heavy chain V r
15	27	30.0	12	2 S25056	Ig heavy chain - m
16	27	30.0	13	2 S47365	T-cell antigen rec
17	27	30.0	13	2 PH0787	T-cell receptor al
18	27	30.0	14	2 PH1598	Ig H chain V-D-J r
19	27	30.0	23	2 PH1364	Ig heavy chain DJ
20	27	30.0	23	2 PH1723	Ig heavy chain V r
21	27	30.0	23	2 C49732	NADH2 dehydrogenas
22	27	30.0	25	2 PH1733	Ig heavy chain V r
23	27	30.0	25	2 A60286	heat-stable serine
24	26.5	29.4	15	2 PH0770	T-cell receptor be
25	26	28.9	12	2 PH1463	T-cell receptor be
26	26	28.9	12	2 PH1470	T-cell receptor be
27	26	28.9	12	2 PH1469	T-cell receptor be
28	26	28.9	12	2 PH1457	T-cell receptor be
29	26	28.9	13	2 S47372	T-cell antigen rec

30 26 28.9 16 2 E53284 T-cell receptor be
31 26 28.9 17 2 B49255 T-cell receptor be
32 26 28.9 17 2 C49255 T-cell receptor be
33 26 28.9 21 2 PH1730 Ig heavy chain V r
34 26 28.9 23 2 PH1707 Ig heavy chain V r
35 26 28.9 23 2 PH1722 Ig heavy chain V r
36 26 28.9 23 2 PH1724 Ig heavy chain V r
37 26 28.9 23 2 PH1725 Ig heavy chain V r
38 26 28.9 23 2 PH1727 Ig heavy chain V r
39 26 28.9 24 2 PH1710 Ig heavy chain V r
40 26 28.9 24 2 PH1712 Ig heavy chain V r
41 26 28.9 24 2 PH1713 Ig heavy chain V r
42 26 28.9 24 2 PH1916 T-cell receptor be
43 25 27.8 12 2 S26552 T-cell receptor be
44 25 27.8 12 2 PH1462 T-cell receptor be
45 25 27.8 12 2 PH1461 T-cell receptor be

ALIGNMENTS

RESULT 1

A49038

Ig lambda chain V region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C:Accession: A49038

R:Weiss, U.; Zobebelein, R.; Rajewsky, K.

Eur. J. Immunol. 22, 511-517, 1992

A:Title: Accumulation of somatic mutants in the B cell compartment after primary immu

A:Reference number: A49038; MUID:92164733; PMID:1537385

A:Accession: A49038

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-25 <WEI>

A:Cross-references: GB:S85732; NID:g246314; PIDN:ANB21559.1; PID:g246315

A:Experimental source: spleen

A>Note: sequence extracted from NCBI backbone (NCBIN:85732, NCBIIP:85757)

C:Keywords: heterotetramer; immunoglobulin

Query Match 41.1%; Score 37; DB 2; Length 25;

Best Local Similarity 33.3%; Pred No. 4.8; Mismatches 5; Gaps 0;

Matches 6; Conservative 5; Indels 7; Gaps 0;

QY 1 VSKLEASTIRQYYFGDA 18

II : : : II : :

Db 1 VSMWKLGLTMERYYYGSS 18

RESULT 2

C57001

endo-1,4-beta-xylanase (EC 3.2.1.8) 4 - Streptomyces roseiscleroticus (fragment)

C:Species: Streptomyces roseiscleroticus

C>Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 23-May-1997

C:Accession: C57001

R:Grabski, A.C.; Forrester, I.T.; Patel, R.; Jeffries, T.W.

Protein Expr. Purif. 4, 120-129, 1993

A:Title: Characterization and N-terminal amino acid sequences of beta-(1-4)endoxylan-

A:Reference number: A57001; MUID:93229899; PMID:8471845

A:Accession: C57001

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-25 <GRA>

A:Experimental source: strain NRRL B-11019

A>Note: sequence extracted from NCBI backbone (NCBIIP:130009)

C:Function:

A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans

A:Pathway: xylan degradation

C:Superfamily: Streptomyces endo-1,4-beta-xylanase A; Streptomyces endo-1,4-beta-xyli-

C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation

Query Match 38.9%; Score 35; DB 2; Length 25;

Best Local Similarity 47.1%; Pred. No. 11;

Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 2 SKLEASTIRQYFYFGDA 18
| | | : | | | |
Db 3 STLGAARQSGYFYFGTA 19

RESULT 3
PH0746

T-cell receptor beta chain (B28) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
C:Accession: PH0746
R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A:Title: T cell receptor genes in a series of class I major histocompatibility complex-I allelic exclusion and antigen-specific repertoire.
A:Reference number: PH0746; MUID:92078846; PMID:1836010
A:Accession: PH0746
A:Molecule type: mRNA
A:Residues: 1-12 <CAS>
A:Cross-references: EMBL:X60837; NID:g50098; PIDN:CAA43230.1; PID:g50099
A:Experimental source: T lymphocyte
C:Keywords: T-cell receptor

Query Match 37.8%; Score 34; DB 2; Length 12;
Best Local Similarity 63.6%; Pred. No. 7.3;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 6 ASTIRQYFYFG 16
| | | : | | | |
Db 2 ASSRQYFYFG 12

RESULT 4
PH1466

T-cell receptor beta chain (clone A3/74.1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 11-Apr-1995
C:Accession: PH1466
R:Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; Kourilsky, J. Exp. Med. 177, 811-820, 1993
A:Title: T cell receptor selection by and recognition of two class I major histocompatibility complex-I allelic exclusion and antigen-specific repertoire.
A:Reference number: PH1430; MUID:93171821; PMID:8436911
A:Accession: PH1466
A:Molecule type: mRNA
A:Residues: 1-12 <CAS>
A:Experimental source: cytolytic T-lymphocyte
C:Superfamily: immunoglobulin homology
C:Keywords: receptor; T-cell

Query Match 36.7%; Score 33; DB 2; Length 12;
Best Local Similarity 54.5%; Pred. No. 11;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ASTIRQYFYFG 16
| | | : | | | |
Db 2 ASSRQYFYFG 12

RESULT 5
PH0771

T-cell receptor beta chain (PE5.1.1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
C:Accession: PH0771
R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A:Title: T cell receptor genes in a series of class I major histocompatibility complex-I allelic exclusion and antigen-specific repertoire.
A:Reference number: PH0746; MUID:92078846; PMID:1836010
A:Accession: PH0771
A:Molecule type: mRNA

A:Residues: 1-12 <CAS>
A:Cross-references: EMBL:X60855; NID:g53624; PIDN:CAA43255.1; PID:g53625
A:Experimental source: T lymphocyte
C:Keywords: T-cell receptor

Query Match 35.6%; Score 32; DB 2; Length 12;
Best Local Similarity 54.5%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 6 ASTIRQYFYFG 16
| | | : | | | |
Db 2 ASSRQYFYFG 12

RESULT 6
S26559

T-cell receptor beta chain (clone Cw3/Cas15) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999
C:Accession: S26559
R:Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Gournier, H.; Barra, C.; J. Exp. Med. 176, 439-447, 1992
A:Title: H-2-restricted cytolytic T lymphocytes specific for HLA display T cell receptor I allelic exclusion and antigen-specific repertoire.
A:Reference number: S26512; MUID:92364546; PMID:1380061
A:Accession: S26559
A:Molecule type: mRNA
A:Residues: 1-12 <CAS>
A:Cross-references: EMBL:X68009
A:Experimental source: cytolytic T-lymphocyte, clone Cw3/Cas15
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match 34.4%; Score 31; DB 2; Length 12;
Best Local Similarity 54.5%; Pred. No. 26;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 6 ASTIRQYFYFG 16
| | | : | | | |
Db 2 ASSRQYFYFG 12

RESULT 7
B49480

major immunophilin hsp56 - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 04-Sep-1998
C:Accession: B49480
R:Yem, A.W.; Reardon, I.M.; Leone, J.W.; Heinrichson, R.L.; Delbel Jr., M.R. Biochemistry 32, 12571-12576, 1993
A:Title: An active FK506-binding domain of 17,000 daltons is isolated following limit digestion of chicken immunophilin hsp56.
A:Reference number: A49480; MUID:94072550; PMID:7504525
A:Accession: B49480
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-24 <YEM>
A:Experimental source: thymus
A:Note: sequence extracted from NCBI backbone (NCBIP:142439)
C:Superfamily: human FK506-binding protein FKBP1; BKBP-type peptidylprolyl isomerase

Query Match 33.3%; Score 30; DB 2; Length 24;
Best Local Similarity 38.9%; Pred. No. 86;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 1 VSKLEASTIRQYFYFGDA 18
| | | : | | | |
Db 2 VGLCRITCKPEYAYGSA 19

RESULT 8
PH1716

Ig heavy chain V region (clone ASC-14) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999

C:Accession: PH1716
 R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
 J. Exp. Med. 178, 295-307, 1993
 A:Title: Antigen-driven B cell differentiation in vivo.
 A:Reference number: PH1675; MUID:93301607; PMID:8315385
 A:Accession: PH1716
 A:Molecule type: mRNA
 A:Residues: 1-25 <MCH>
 A:Experimental source: B cell
 A:Note: the authors translated the codon GFA for residue 11 as Thr and ACA for residue 1
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotrimer; immunoglobulin

Query Match 33.3%; Score 30; DB 2; Length 25;

Best Local Similarity 36.4%; Pred. No. 91;

Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 TIRQYFYFGDA 18

Db 10 SVETSYYGSA 20

RESULT 9

S26549
 T-cell receptor beta chain (clone Cw3/A8, Cw3/Cas1) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 13-Jan-1995 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999

C:Accession: S26549; S26550

R:Casanova, J.L.; Gerotini, J.C.; Matthes, M.; Necker, A.; Gournier, H.; Barra, C.; Wido

J. Exp. Med. 176, 439-447, 1992

A:Title: H-2-restricted cytolytic T lymphocytes specific for HLA display T cell receptor

A:Reference number: S26512; MUID:92364546; PMID:1380061

A:Accession: S26549

A:Molecule type: mRNA

A:Residues: 1-12 <CAS>

A:Cross-references: EMBL:X67999

A:Experimental source: cytolytic T-lymphocyte, clone Cw3/A8

A:Accession: S26550

A:Molecule type: mRNA

A:Residues: 1-12 <CA2>

A:Cross-references: EMBL:X68000

A:Experimental source: cytolytic T-lymphocyte, clone Cw3/Cas1

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: T-cell receptor

Query Match 31.1%; Score 28; DB 2; Length 12;

Best Local Similarity 45.5%; Pred. No. 90;

Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 ASTIRQYFYFG 16

Db 2 ASSLGTLVFG 12

RESULT 10

PH1454

T-cell receptor alpha chain (clone A3/72.2) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 11-Apr-1995

C:Accession: PH1454

R:Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; Ko

J. Exp. Med. 177, 811-820, 1993

A:Title: T cell receptor selection by and recognition of two class I major histocompatib

A:Reference number: PH1430; MUID:93171821; PMID:8436911

A:Accession: PH1454

A:Molecule type: mRNA

A:Residues: 1-12 <CAS>

A:Experimental source: cytolytic T-lymphocyte

C:Superfamily: immunoglobulin homology

C:Keywords: receptor; T-cell

Query Match 31.1%; Score 28; DB 2; Length 12;

Best Local Similarity 54.5%; Pred. No. 90;

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 ASTIRQYFYFG 16

Db 2 AVTGYQNFYFG 12

RESULT 11

PH1368

Ig heavy chain DJ region (clone C111-112) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: PH1368

R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.

J. Exp. Med. 176, 1577-1581, 1992

A:Title: Predominance of fetal type DJH joining in young children with B precursor ly

A:Reference number: PH1302; MUID:93094761; PMID:1460419

A:Accession: PH1368

A:Molecule type: DNA

A:Residues: 1-18 <WAS>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotrimer; immunoglobulin

Query Match 31.1%; Score 28; DB 2; Length 18;

Best Local Similarity 80.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYVFG 16

Db 6 EYVFG 10

RESULT 12

PH1629

Ig H chain V-D-J region (clone B-less 155) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999

C:Accession: PH1629

R:Levinson, D.A.; Campos-Torres, J.; Leder, P.

J. Exp. Med. 178, 317-329, 1993

A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less m

A:Reference number: PH1580; MUID:93301609; PMID:8315387

A:Accession: PH1629

A:Molecule type: DNA

A:Residues: 1-18 <LEV>

A:Experimental source: bone marrow pre-B lymphocyte

C:Keywords: immunoglobulin

Query Match 31.1%; Score 28; DB 2; Length 18;

Best Local Similarity 50.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 9 IRQYFYFG 16

Db 2 VRRDYDG 9

RESULT 13

S04169

IgA-binding protein - Streptococcus sp. (fragment)

N:Alternate names: protein Arp

C:Species: Streptococcus sp.

C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 23-Mar-1993

C:Accession: S04169

R:Lindahl, G.; Akerstroem, B.

Mol. Microbiol. 3, 239-247, 1989

A:Title: Receptor for IgA in group A streptococci: cloning of the gene and characteri

A:Reference number: S04169; MUID:89343628; PMID:2668688

A:Accession: S04169

A:Molecule type: protein

A:Residues: 1-19 <LIN>

Query Match

31.1%; Score 28; DB 2; Length 19;

Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 KLEASTIROE 12
| | | | |
Db 6 KAESTVKAE 15

RESULT 14

PHI734

Ig heavy chain V region (clone GCC-14) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999

C:Accession: PHI734

R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.

J. Exp. Med. 178, 295-307, 1993

A:Title: Antigen-driven B cell differentiation in vivo.

A:Reference number: PHI675; MUID:93301607; PMID:8315385

A:Accession: PHI734

A:Molecule type: mRNA

A:Residues: 1-25 <MCH>

A:Experimental source: B cell

A:Note: the authors translated the codon ACA for residue 13 as Ala

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 31.1%; Score 28; DB 2; Length 25;
Best Local Similarity 44.4%; Pred. No. 2.1e+02;

Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 RQEYVFGDA 18
| | | | |
Db 14 RYSYVYGGSS 22

RESULT 15

S25056

Ig heavy chain - mouse

C:Species: Mus musculus (house mouse)

C:Date: 25-Feb-1994 #sequence_revision 01-Dec-1995 #text_change 20-Jun-2000

C:Accession: S25056

R:Jacob, J.; Kelsoe, G.

submitted to the EMBL Data Library, July 1992

A:Description: In situ studies on the primary immune response to (4-hydroxy-3-nitrophenyl

A:Reference number: S25024

A:Accession: S25056

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-12 <JAC>

A:Cross-references: EMBL:X67386; NID:g50927; PIDN:CAA47798.1; PID:gl333920

C:Superfamily: Immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match 30.0%; Score 27; DB 2; Length 12;
Best Local Similarity 44.4%; Pred. No. 1.4e+02;

Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 RQEYVFGDA 18
| | | | |
Db 2 RYPYVYGGSS 10

Search completed: April 23, 2003, 13:48:56
Job time : 10.4045 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:38:28 ; Search time 8.49438 Seconds
(without alignments)
62.349 Million cell updates/sec

Title: US-09-836-073-18

Perfect score: 90
Sequence: 1 VSKLEASTIRQYFYFGDA 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 147762

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42.5	47.2	18	4	US-09-316-630-3
2	42.5	47.2	18	4	US-09-316-630-4
3	36	40.0	12	1	US-07-894-212A-4
4	36	40.0	12	1	US-07-893-928A-2
5	35	38.9	25	1	US-08-453-289-4
6	35	38.9	25	2	US-08-574-086-4
7	31	34.4	20	1	US-07-748-344B-7
8	31	34.4	20	2	US-08-954-985A-7
9	31	34.4	20	4	US-08-137-086-7
10	31	34.4	21	1	US-07-748-344B-8
11	31	34.4	21	2	US-08-954-985A-8
12	31	34.4	21	4	US-08-137-086-8
13	29	33.2	14	1	US-08-475-213-7
14	28.5	31.7	10	2	US-08-618-696-11
15	28.5	31.7	10	3	US-09-033-753-11
16	28.5	31.7	11	2	US-08-618-696-7
17	28.5	31.7	11	2	US-08-618-696-20
18	28.5	31.7	11	3	US-09-033-753-7
19	28.5	31.7	11	3	US-09-033-753-2
20	28	31.1	10	2	US-08-618-696-2
21	28	31.1	10	2	US-08-618-696-10
22	28	31.1	10	3	US-09-033-753-2
23	28	31.1	10	3	US-09-033-753-10
24	28	31.1	11	2	US-08-618-696-1
25	28	31.1	11	2	US-08-618-696-6
26	28	31.1	11	2	US-08-618-696-16
27	28	31.1	11	2	US-08-618-696-19

28	28	31.1	11	3	US-09-033-753-1	Sequence 1, Appli
29	28	31.1	11	3	US-09-033-753-6	Sequence 6, Appli
30	28	31.1	11	3	US-09-033-753-16	Sequence 16, Appli
31	28	31.1	11	3	US-09-033-753-19	Sequence 19, Appli
32	28	31.1	17	4	US-08-990-823-87	Sequence 87, Appli
33	27.5	30.6	23	3	US-08-963-121C-1	Sequence 1, Appli
34	27.5	30.6	23	4	US-09-543-513-1	Sequence 1, Appli
35	27.5	30.6	23	5	PCT-US95-04803-1	Sequence 1, Appli
36	27	30.0	14	4	US-09-106-568E-24	Sequence 24, Appli
37	27	30.0	14	4	US-09-106-568E-61	Sequence 61, Appli
38	27	30.0	14	4	US-09-106-568E-65	Sequence 65, Appli
39	27	30.0	14	4	US-09-106-568E-81	Sequence 81, Appli
40	27	30.0	15	2	US-08-553-257A-66	Sequence 66, Appli
41	27	30.0	15	4	US-09-116-492A-20	Sequence 20, Appli
42	27	30.0	16	1	US-08-453-378-10	Sequence 10, Appli
43	27	30.0	17	1	US-08-453-378-1	Sequence 1, Appli
44	27	30.0	18	1	US-08-176-500-32	Sequence 32, Appli
45	27	30.0	18	1	US-08-471-052A-32	Sequence 32, Appli

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match 47.2%; Score 42.5; DB 4; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.23;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 4 LEASTIRQ-EYFEGD 17
||| |
DB 3 LEAKTCHQIEYFEGD 17

RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

; PRIOR FILING DATE: 1997-10-06
 ; PRIOR APPLICATION NUMBER: 08/321,427
 ; PRIOR FILING DATE: 1994-10-11
 ; PRIOR APPLICATION NUMBER: 60/086,527
 ; PRIOR FILING DATE: 1998-05-22
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
 ; OTHER INFORMATION: this peptide is biotinylated
 US-09-316-630-4

Query Match 47.2%; Score 42.5; DB 4; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.23;
 Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 4 LEASTIRQ-EYFPGD 17
 ||| | |||||
 Db 3 LEAKICHQIEYFPGD 17

RESULT 3
 US-07-894-212A-4
 ; Sequence 4, Application US/07894212A
 ; Patent No. 5366883
 ; GENERAL INFORMATION:
 ; APPLICANT: ASADA, KIYOZO
 ; APPLICANT: UEMORI, TAKASHI
 ; APPLICANT: MUKAI, HIROYUKI
 ; APPLICANT: KATO, IKUNOSHUKI
 ; APPLICANT: LADERMAN, KENNETH
 ; APPLICANT: ANFINSEN, CHRISTIAN
 ; TITLE OF INVENTION: THE ALPHA-AMYLASE GENE
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
 ; STREET: 1100 NEW YORK AVENUE, N.W.
 ; CITY: WASHINGTON, D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/894,212A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: KOKULIS, PAUL N.
 ; REGISTRATION NUMBER: 16773
 ; REFERENCE/DOCKET NUMBER: 95469/C-1195
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 861-3000
 ; TELEFAX: (202) 822-0944
 ; TELEX: 6714627 CUSH
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 12 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-07-894-212A-4

Query Match 40.0%; Score 36; DB 1; Length 12;
 Best Local Similarity 85.7%; Pred. No. 2.2;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 IRQEYF 15
 : |||||
 Db 5 MRQEYF 11
 : |||||
 RESULT 4
 US-07-893-928A-2
 ; Sequence 2, Application US/07893928A
 ; Patent No. 5578479
 ; GENERAL INFORMATION:
 ; APPLICANT: LADERMAN, KENNETH
 ; APPLICANT: ANFINSEN, CHRISTIAN
 ; TITLE OF INVENTION: a-AMYLASE FROM HYPERTHERMOPHILIC
 ; TITLE OF INVENTION: ARCHAEBACTERIUM
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
 ; STREET: 1100 NEW YORK AVENUE, N.W.
 ; CITY: WASHINGTON, D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Tape
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/893,928A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: KOKULIS, PAUL N.
 ; REGISTRATION NUMBER: 16773
 ; REFERENCE/DOCKET NUMBER: 95470/C-1197
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 861-3000
 ; TELEFAX: (202) 822-0944
 ; TELEX: 6714627 CUSH
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 12 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-07-893-928A-2

Query Match 40.0%; Score 36; DB 1; Length 12;
 Best Local Similarity 85.7%; Pred. No. 2.2;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 IRQEYF 15
 : |||||
 Db 5 MRQEYF 11

RESULT 5
 US-08-453-289-4
 ; Sequence 4, Application US/08453289
 ; Patent No. 5498534
 ; GENERAL INFORMATION:
 ; APPLICANT: Jeffries, Thomas W
 ; APPLICANT: Grabski, Anthony C
 ; APPLICANT: Patel, Rajesh N
 ; APPLICANT: Elegir, Graziano
 ; APPLICANT: Szakacs, George
 ; TITLE OF INVENTION: Method of Removing Color from Kraft Wood
 ; TITLE OF INVENTION: Pulp
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Janet I. Stockhausen
 ; STREET: One Gifford Pinchot Drive

CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53705
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,289
FILING DATE:
FILING DATE: 25-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Stockhausen, Janet I
REGISTRATION NUMBER: 34,256
TELEPHONE: 608-231-9504
TELEFAX: 608-231-9508
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptomyces roseiscleroticus
STRAIN: NRRLB-11019
US-08-453-289-4

Query Match 38.9%; Score 35; DB 1; Length 25;
Best Local Similarity 47.1%; Pred. No. 7.8;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 SKLEASTIRQYFVGD 18
| | | : | | | |
Db 3 STLGAQAQSGYFVGT 19

RESULT 6
US-08-574-086-4
Sequence 4, Application US/08574086
Patent No. 5834301
GENERAL INFORMATION:
APPLICANT: Jeffries, Thomas W
APPLICANT: Grabski, Anthony C
APPLICANT: Patel, Rajesh N
APPLICANT: Elegir, Graziano
APPLICANT: Szakacs, George
TITLE OF INVENTION: Method of Removing Color from Kraft Wood
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janet I. Stockhausen
STREET: One Gifford Pinchot Drive
CITY: Madison
STATE: Wisconsin
COUNTRY: USA
ZIP: 53705
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/574,086
FILING DATE: 18-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/453,289
FILING DATE:
APPLICATION NUMBER: US/08/257,965
FILING DATE:
APPLICATION NUMBER: US 07/857,060
FILING DATE: 25-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Stockhausen, Janet I
REGISTRATION NUMBER: 34,256
TELEPHONE: 608-231-9504
TELEFAX: 608-231-9508
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptomyces roseiscleroticus
STRAIN: NRRLB-11019
US-08-574-086-4

Query Match 38.9%; Score 35; DB 2; Length 25;
Best Local Similarity 47.1%; Pred. No. 7.8;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 SKLEASTIRQYFVGD 18
| | | : | | | |
Db 3 STLGAQAQSGYFVGT 19

RESULT 7
US-07-748-344B-7
Sequence 7, Application US/07748344B
Patent No. 5227154
GENERAL INFORMATION:
APPLICANT: REYNOLDS, ERIC CHARLES
TITLE OF INVENTION: PHOSPHOPEPTIDES FOR THE
TREATMENT OF DENTAL CALCULUS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: ANDRUS, SCALES, STARKE & SAWALL
STREET: 100 EAST WISCONSIN AVE., SUITE 1100
CITY: MILWAUKEE
STATE: WISCONSIN
COUNTRY: USA
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/748,344B
FILING DATE: 19910822
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: SARA, CHARLES S
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: C.8493-87
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 255-2022


```
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: Protein
;; FEATURE:
;; FEATURE: 7 Phosphoserine
;; FEATURE:
;; FEATURE:
;; FEATURE: Post-translationally phosphorylated serine
;; FEATURE:
;; FEATURE: Phosphoserine
;; FEATURE: 8
;; FEATURE:
;; FEATURE: Post-translationally phosphorylated serine
;; FEATURE:
;; FEATURE: Phosphoserine
;; FEATURE: 9
;; FEATURE:
;; FEATURE: Post-translationally phosphorylated serine
;; FEATURE:
;; FEATURE: Phosphoserine
;; FEATURE: 15
;; FEATURE:
;; FEATURE: Post-translationally phosphorylated serine
;; FEATURE:
;; FEATURE:
US-08-137-086-7

Query Match 34.4%; Score 31; DB 4; Length 20;
Best Local Similarity 57.1%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 VSKLEASTIRQEYY 14
  || | | | | |
Db 6 VSSSEESIISQETY 19
```

```
RESULT 10
US-07-748-344B-8
; Sequence 8, Application US/07748344B
; Patent No. 5227154
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES
; TITLE OF INVENTION: PHOSPHOPEPTIDES FOR THE
; TREATMENT OF DENTAL CALCULUS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDRUS, SCEALES, STARKE & SAWALL
; STREET: 100 EAST WISCONSIN AVE., SUITE 1100
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/748,344B
; FILING DATE: 19910822
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: SARA, CHARLES S
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: C.8493-87
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 255-2022
; TELEFAX: (608) 255-2182
; TELEX: 26832 ANDSTARK
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
```

```
;; TYPE: AMINO ACID
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: Protein
;; FEATURE:
;; FEATURE: 8 Phosphoserine
;; FEATURE:
;; FEATURE: Post-translationally phosphorylated serine
;; FEATURE:
;; FEATURE: Phosphoserine
;; FEATURE: 9
;; FEATURE:
;; FEATURE: Post-translationally phosphorylated serine
;; FEATURE:
;; FEATURE: Phosphoserine
;; FEATURE: 10
;; FEATURE:
;; FEATURE: Post-translationally phosphorylated serine
;; FEATURE:
;; FEATURE: Phosphoserine
;; FEATURE: 16
;; FEATURE:
;; FEATURE: Post-translationally phosphorylated serine
US-07-748-344B-8

Query Match 34.4%; Score 31; DB 1; Length 21;
Best Local Similarity 57.1%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 VSKLEASTIRQEYY 14
  || | | | | |
Db 7 VSSSEESIISQETY 20

RESULT 11
US-08-954-985A-8
; Sequence 8, Application US/08954985A
; Patent No. 5981475
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES
; TITLE OF INVENTION: A Treatment for Sensitive Teeth
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dental School, The University of Melbourne
; STREET: 711 Elizabeth Street
; CITY: Melbourne
; STATE: Victoria
; COUNTRY: Australia
; ZIP: 3000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM
; OPERATING SYSTEM: Windows 95
; SOFTWARE: Ascii
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,985A
; FILING DATE: 21-OCT-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,479
; FILING DATE: 22-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Wozny, Thomas M
; REGISTRATION NUMBER: 28,922
; REFERENCE/DOCKET NUMBER: 322-00033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 271-7590
; TELEFAX: (414) 271-5770
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: Amino Acid
```

```

; TOPOLOGY: Linear
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 8
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 9
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 10
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 16
; OTHER INFORMATION: Post-translationally phosphorylated serine
; FEATURE:
; NAME/KEY: Phosphoserine
; LOCATION: 20
; OTHER INFORMATION: Post-translationally phosphorylated serine
; US-08-954-985A-8

```

Query Match 34.4%; Score 31; DB 2; Length 21;
 Best Local Similarity 57.1%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 VSKLEASTIRQY 14
|| | | | | | |
Db 7 VSSSESIISQY 20

```

```

RESULT 12
US-08-137-086-8
; Sequence 8, Application US/08137086
; Patent No. 6448374
; GENERAL INFORMATION:
; APPLICANT: REYNOLDS, ERIC CHARLES
; TITLE OF INVENTION: PRODUCTION OF PHOSPHOPEPTIDES FROM CASEIN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDRUS, SCALES, STARKE & SAWALL
; STREET: 100 EAST WISCONSIN AVE., SUITE 1100
; CITY: MILWAUKEE
; STATE: WISCONSIN
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/137,086
; FILING DATE: 04-MAR-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: SARA, CHARLES S
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: C.8493-87
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 255-2022
; TELEFAX: (608) 255-2182
; TELEX: 26832 ANDSTARK
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: Amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; FEATURE:
; FEATURE: Phosphoserine

```

```

; FEATURE: 8
; FEATURE: Post-translationally phosphorylated serine
; FEATURE: Phosphoserine
; FEATURE: 9
; FEATURE: Post-translationally phosphorylated serine
; FEATURE: Phosphoserine
; FEATURE: 10
; FEATURE: Post-translationally phosphorylated serine
; FEATURE: Phosphoserine
; FEATURE: 16
; FEATURE: Post-translationally phosphorylated serine
; FEATURE: 20
; US-08-137-086-8

```

Query Match 34.4%; Score 31; DB 4; Length 21;
 Best Local Similarity 57.1%; Pred. No. 33;
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 VSKLEASTIRQY 14
|| | | | | | |
Db 7 VSSSESIISQY 20

```

```

RESULT 13
US-08-475-213-7
; Sequence 7, Application US/08475213
; Patent No. 5783674
; GENERAL INFORMATION:
; APPLICANT: Geysen, Hendrik M.
; TITLE OF INVENTION: Method for the use and Synthesis of
; TITLE OF INVENTION: Peptides
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,213
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/752,437
; FILING DATE: 06-SEP-1991
; APPLICATION NUMBER: WO pct/au90/00062
; FILING DATE: 16-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ2788/89
; FILING DATE: 17-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31259
; REFERENCE/DOCKET NUMBER: 0240.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids

```

;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-475-213-7

Query Match 32.2%; Score 29; DB 1; Length 14;
Best Local Similarity 87.5%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 SKLEASTI 9
Db 6 SKLPASTI 13

RESULT 14
US-08-618-696-11
; Sequence 11, Application US/08618696
; Patent No. 5861475
; GENERAL INFORMATION:
; APPLICANT: COOPER, JR., J. ALLEN D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; INHIBITION OF PHAGOCYTES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,696
; FILING DATE: 20-MAR-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/995,269
; FILING DATE: 12/21/92
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UOAB:002/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-320-7200
; TELEFAX: 512-474-7577
; TELEX: NOT APPLICABLE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-618-696-11

Query Match 31.7%; Score 28.5; DB 2; Length 10;
Best Local Similarity 53.8%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 5 EASTIRQEVYFGD 17
Db 1 EAS-----YFFGD 8

RESULT 15
US-09-033-753-11
; Sequence 11, Application US/09033753
; Patent No. 6017863
; GENERAL INFORMATION:

;; APPLICANT: COOPER, JR., J. ALLEN D.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;; INHIBITION OF PHAGOCYTES
;; NUMBER OF SEQUENCES: 21
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ARNOLD, WHITE & DURKEE
;; STREET: P.O. BOX 4433
;; CITY: HOUSTON
;; STATE: TEXAS
;; COUNTRY: USA
;; ZIP: 77210
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: FLOPPY DISK
;; COMPUTER: IBM PC COMPATIBLE
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WORDPERFECT 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/033,753
;; FILING DATE:
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/618,696
;; FILING DATE: 20-MAR-1996
;; APPLICATION NUMBER: 07/995,269
;; FILING DATE: 12/21/92
;; ATTORNEY/AGENT INFORMATION:
;; NAME: PARKER, DAVID L.
;; REGISTRATION NUMBER: 32,165
;; REFERENCE/DOCKET NUMBER: UOAB:002/PAR
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 512-320-7200
;; TELEFAX: 512-474-7577
;; TELEX: NOT APPLICABLE
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10 amino acid residues
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-09-033-753-11

Query Match 31.7%; Score 28.5; DB 3; Length 10;
Best Local Similarity 53.8%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 0; Indels 5; Gaps 1;

QY 5 EASTIRQEVYFGD 17
Db 1 EAS-----YFFGD 8

Search completed: April 23, 2003, 13:50:28
Job time: 8.49438 secs

GenCore version 5.1.4.p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:43:09 ; Search time 9.91011 Seconds
(without alignments)
145.542 Million cell updates/sec

Title: US-09-836-073-18

Perfect score: 90
Sequence: 1 VSKLEASTIRQYFYFGDA 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 61792

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	90	100.0	18	9 US-09-836-073-18	Sequence 18, Appl
2	44.5	49.4	16	9 US-09-836-073-19	Sequence 19, Appl
3	43.5	48.3	18	9 US-09-836-073-2	Sequence 2, Appl
4	42.5	47.2	17	9 US-09-836-073-13	Sequence 13, Appl
5	42.5	47.2	18	9 US-09-836-073-1	Sequence 1, Appl
6	42.5	47.2	18	9 US-09-836-073-9	Sequence 9, Appl
7	42.5	47.2	18	9 US-09-836-073-14	Sequence 14, Appl
8	42	46.7	19	9 US-09-836-073-16	Sequence 16, Appl
9	38.5	42.8	18	9 US-09-836-073-4	Sequence 4, Appl
10	37	41.1	18	9 US-09-836-073-5	Sequence 5, Appl
11	37	41.1	18	9 US-09-836-073-15	Sequence 15, Appl
12	34.5	38.3	18	9 US-09-836-073-11	Sequence 11, Appl
13	34.5	38.3	18	9 US-09-836-073-12	Sequence 12, Appl
14	33.5	37.2	18	9 US-09-836-073-7	Sequence 7, Appl
15	33.5	37.2	18	9 US-09-836-073-10	Sequence 10, Appl
16	32.5	36.1	20	9 US-09-836-073-17	Sequence 17, Appl
17	32	35.6	20	9 US-10-078-622-597	Sequence 597, App
18	30.5	33.9	18	9 US-09-836-073-3	Sequence 3, Appl
19	29	32.2	15	9 US-09-880-748-3083	Sequence 3083, Ap

Sequence 77, Appl
Sequence 77, Appl
Sequence 2, Appl
Sequence 87, Appl
Sequence 87, Appl
Sequence 87, Appl
Sequence 3145, Ap
Sequence 181, App
Sequence 3010, Ap
Sequence 24, Appl
Sequence 61, Appl
Sequence 65, Appl
Sequence 81, Appl
Sequence 48, Appl
Sequence 155, App
Sequence 156, App
Sequence 179, App
Sequence 49, Appl
Sequence 59, Appl
Sequence 49, Appl
Sequence 59, Appl
Sequence 32, Appl
Sequence 32, Appl
Sequence 8, Appl
Sequence 17, Appl

20 9 US-10-044-703-77
21 10 US-09-813-333-77
22 15 9 US-09-273-217-2
23 17 9 US-09-996-634-87
24 17 9 US-09-997-181-87
25 17 9 US-09-997-182-87
26 17 9 US-09-997-182-87
27 19 9 US-09-880-748-3145
28 21 9 US-10-062-710-181
29 21 9 US-09-880-748-3010
30 14 9 US-09-994-595-24
31 14 9 US-09-994-595-61
32 14 9 US-09-994-595-65
33 14 9 US-09-994-595-81
34 15 9 US-10-153-159-48
35 15 9 US-10-153-176-48
36 16 9 US-10-174-105A-155
37 16 9 US-10-174-105A-156
38 16 9 US-10-174-105A-179
39 19 9 US-10-153-159-49
40 19 9 US-10-153-159-59
41 19 9 US-10-153-176-49
42 20 9 US-10-153-176-59
43 20 9 US-10-042-945-32
44 20 10 US-09-757-417-32
45 26.5 29.4 18 9 US-09-836-073-8
26 12 10 US-09-564-329A-17

ALIGNMENTS

RESULT 1
US-09-836-073-18
; Sequence 18, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Mosquito
US-09-836-073-18

Query Match 100.0%; Score 90; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.7e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSKLEASTIRQYFYFGDA 18

Db 1 VSKLEASTIRQYFYFGDA 18

RESULT 2

US-09-836-073-19
; Sequence 19, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Drosophila
US-09-836-073-19

Query Match 49.4%; Score 44.5; DB 9; Length 16;
Best Local Similarity 71.4%; Pred. No. 0.23;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 5 EASTIRQ-EYFQD 17
| : ||| |||||
Db 2 ERALRQVEYFQD 15

RESULT 3

; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match 48.3%; Score 43.5; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.39;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 4 LEASTIRQ-EYFQD 17
| : ||| |||||
Db 3 LEAICQIEYFQD 17

RESULT 4

; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match 47.2%; Score 42.5; DB 9; Length 17;
Best Local Similarity 66.7%; Pred. No. 0.55;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 4 LEASTIRQ-EYFQD 17
| : ||| |||||
Db 2 LEAKICQIEYFQD 16

RESULT 5

US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match 47.2%; Score 42.5; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.59;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 4 LEASTIRQ-EYFQD 17
| : ||| |||||
Db 3 LEAKICQIEYFQD 17

RESULT 6

US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match 47.2%; Score 42.5; DB 9; Length 18;
Best Local Similarity 66.7%; Pred. No. 0.59;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 4 LEASTIRQ-EYFQD 17
| : ||| |||||
Db 3 LEAKICQIEYFQD 17

RESULT 7

US-09-836-073-14

```
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14

Query Match
Best Local Similarity 47.2%; Score 42.5; DB 9; Length 18;
Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 4 LEASTIRQ-EYYFGD 17
   ||| | |||||
Db 3 LEAKICHQIEYFGD 17

RESULT 8
US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match
Best Local Similarity 46.7%; Score 42; DB 9; Length 19;
Matches 10; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

Qy 4 LEASTIRQ-EYYFGD 17
   ||| | |||||
Db 3 LEAKICHQIEYFGD 18

RESULT 9
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Xenopus
US-09-836-073-15

Query Match
Best Local Similarity 41.1%; Score 37; DB 9; Length 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 EYFQGD 17
   |||||
Db 12 EYFQGD 17

RESULT 11
US-09-836-073-15
; Sequence 15, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Xenopus
US-09-836-073-15

Query Match
Best Local Similarity 41.1%; Score 37; DB 9; Length 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 EYFQGD 17
   |||||
Db 12 EYFQGD 17
```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFQD 17
 |||||
 Db 12 EYFQD 17

RESULT 12

US-09-836-073-11

; Sequence 11, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-11

Query Match 38.3%; Score 34.5; DB 9; Length 18;
 Best Local Similarity 60.0%; Pred. No. 14;
 Matches 9; Conservative 0; Mismatches 5; Indels 5; Gaps 1;

QY 4 LEASTIRQ-EYFQD 17
 |||||
 Db 3 LEAKICHQIEYQD 17

RESULT 13

US-09-836-073-12

; Sequence 12, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-12

Query Match 38.3%; Score 34.5; DB 9; Length 18;
 Best Local Similarity 60.0%; Pred. No. 14;
 Matches 9; Conservative 0; Mismatches 5; Indels 5; Gaps 1;

QY 4 LEASTIRQ-EYFQD 17
 |||||
 Db 3 LEAKICHQIEYQD 17

RESULT 14

US-09-836-073-7

; Sequence 7, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-7

Query Match 37.2%; Score 33.5; DB 9; Length 18;
 Best Local Similarity 60.0%; Pred. No. 21;
 Matches 9; Conservative 0; Mismatches 5; Indels 5; Gaps 1;

QY 4 LEASTIRQ-EYFQD 17
 |||||
 Db 3 LEAKICHQIEYQD 17

RESULT 15

US-09-836-073-10

; Sequence 10, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-10

Query Match 37.2%; Score 33.5; DB 9; Length 18;
 Best Local Similarity 60.0%; Pred. No. 21;
 Matches 9; Conservative 0; Mismatches 5; Indels 5; Gaps 1;

QY 4 LEASTIRQ-EYFQD 17
 |||||
 Db 3 LEAKICHQIEYQD 17

Search completed: April 23, 2003, 13:52:11
 Job time : 9.91011 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:28:18 ; Search time 27.2022 Seconds
(without alignments)
88.173 Million cell updates/sec

Title: us-09-836-073-18
Perfect score: 90
Sequence: 1 VSKLEASTIRQEYFGDA 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 320064

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*				
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*			
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*			
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*			
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*			
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*			
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*			
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*			
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*			
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*			
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*			
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*			
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*			
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*			
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*			
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*			
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*			
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*			
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*			
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*			
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*			
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*			
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*			
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Match	Length	Description
1	42.5	47.2	18	21 AAY52200 Human la autoantigen
2	37	41.1	21	14 AAR43394 La/Ssb epitope 17.
3	36	40.0	12	15 AAR46559 Alpha amylase tryp
4	36	40.0	12	15 AAR53628 Pyrococcus furiosu
5	36	40.0	14	21 AAY66883 T cell antigen rec
6	35	38.9	25	17 AAR92123 Streptomyces rosei
7	35	38.9	25	20 AAW80367 N-terminal sequenc
8	34	37.8	20	17 AAR88526 Varicella zoster g
9	33	36.7	22	17 AAW00599 Heparan sulphate p
10	32	35.6	20	21 AAY66945 T cell antigen rec

11	32	35.6	25	23	ABG62324	Eubacterial DNA po
12	31	34.4	9	23	ABP47562	N. meningitidis LO
13	31	34.4	11	23	ABP47702	N. meningitidis LO
14	31	34.4	19	23	ABG62293	Eubacterial DNA po
15	31	34.4	20	8	AA71322	Phosphopeptide 3.
16	31	34.4	20	12	AAR14449	Phosphopeptide #3
17	31	34.4	20	13	AAR28431	Anticariogenic pho
18	31	34.4	20	14	AAR32933	Casin phosphopept
19	31	34.4	20	14	AAR32310	Phosphopeptide der
20	31	34.4	20	14	AAR31239	Phosphopeptide 3.
21	31	34.4	20	15	AAR47820	Sequence of casein
22	31	34.4	20	16	AAR68938	Sodium caseinate t
23	31	34.4	20	21	AAR32802	Phosphopeptide T3
24	31	34.4	21	14	AAR32934	Casin phosphopept
25	31	34.4	21	15	AAR47821	Sequence of casein
26	31	34.4	21	19	AAW66602	Bos alpha-s2-casei
27	31	34.4	24	14	AAR42890	Anti-acid peptide.
28	30	33.3	9	20	AAV42082	Rheumatoid arthrit
29	30	33.3	19	21	AAS52201	Human anti-HBs ant
30	30	33.3	21	23	AAU88609	Insulin/insulin-l
31	29	32.2	10	22	AAU883596	Arabidopsis thalia
32	29	32.2	15	23	ABP47072	Human Blys binding
33	29	32.2	20	18	AAW34602	Partial sequence o
34	29	32.2	20	22	AAEL2952	Mycobacterium tube
35	29	32.2	24	16	AAR65045	Random biotinylati
36	29	32.2	24	20	AAV12811	Human 5' ESR secre
37	29	32.2	25	22	AAU93683	Anti-Rh(D) heavy c
38	29	32.2	25	23	ABG62323	Eubacterial DNA po
39	28.5	31.7	10	15	AAR56297	Synthetic modified
40	28.5	31.7	11	15	AAR56307	Modified Influenza
41	28.5	31.7	11	15	AAR56292	Synthetic derivati
42	28.5	31.7	19	10	AAU90453	Epitope recognised
43	28	31.1	10	15	AAR56293	Native human neutr
44	28	31.1	10	15	AAR56296	Synthetic modified
45	28	31.1	10	21	AAY66804	T cell antigen rec

ALIGNMENTS

RESULT 1
AAY52200
ID AAY52200 standard; peptide; 18 AA.

XX AC AAY52200;

XX DT 14-MAR-2000 (first entry)

XX DE Human la autoantigen peptide (LAP).

XX KW La autoantigen; LAP: Internal ribosome entry site; IRES; translation;
viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
coronavirus; hepatitis virus; rhabdovirus; adenovirus; coxsackie virus;
parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
vesicular stomatitis virus.

XX OS Homo sapiens.

XX PN WO9961613-A2.

XX PD 02-DEC-1999.

XX PF 21-MAY-1999; 99WO-US11281.

XX PR 22-MAY-1998; 98US-0086527.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Das S, Dasgupta A;

XX DR WPI; 2000-062712/05.

XX XX

PT New yeast inhibitory peptide useful for inhibiting viral protein
 PT translation and replication -
 XX Claim 5; Page 57; 81pp; English.
 XX
 XX This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (IRNA see A4245200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rhabdovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.
 XX
 XX Sequence 18 AA;

Query Match 47.2%; Score 42.5; DB 21; Length 18;
 Best Local Similarity 66.7%; Pred. No. 0.9;
 Matches 10; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Oy 4 LEASTIRO-EYVFGD 17
 |||||
 Db 3 LEAKICHOIEYVFGD 17

RESULT 2
 AAR43394
 ID AAR43394 standard; peptide; 21 AA.
 AC AAR43394;

DT 12-MAY-1994 (first entry)
 DE La/SSB epitope 17.

KW Linear; epitope; 60 kD; Ro/SSA; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; RNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B'; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

OS Homo sapiens.

XX WO9321223-A.

XX 28-OCT-1993.

XX 13-APR-1993; 93WO-US03484.

XX 13-APR-1992; 92US-0867819.

XX (OKLA) UNIV OKLAHOMA STATE.

XX Harley JB;

XX WPI; 1993-351658/44.

XX New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating autoimmune disorders e.g. systemic
 PT lupus erythematosus

PS Claim 1; Page 30; 43pp; English.

XX The sequences given in AAR43391-562 are linear epitopes which are

CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (RNP) and the Sm B/B'
 CC polypeptides. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polynucleotide
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.

XX Sequence 21 AA;

Query Match 41.1%; Score 37; DB 14; Length 21;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 12 EYVFGD 17
 |||||
 Db 6 EYVFGD 11

RESULT 3
 AAR46559
 ID AAR46559 standard; Protein; 12 AA.

AC AAR46559;

DT 25-JUN-1994 (first entry)

DE Alpha amylase trypsin and cyanogen bromide fragment.

KW Hyperthermophilic; archaeobacterium; Pyrococcus; P. furiosus; stable;
 KW temperature; liquefaction; starch; amplification.

OS Pyrococcus furiosus.

XX EP579360-A.

XX 19-JAN-1994.

XX 17-MAY-1993; 93EP-0303800.

XX 09-JUN-1992; 92US-0894212.

XX (TAKI) TAKARA SHUZO CO LTD.

XX (UYJO) UNIV JOHNS HOPKINS.

XX Anfinsen CB, Asada K, Kato I, Laderman K, Mukai H;
 XX Uemori T;

XX WPI; 1994-017878/03.

XX New DNA encoding hyper-thermophilic alpha amylase - esp. from
 PT Pyrococcus furiosus,, useful in high temp. starch liquefaction,
 PT and derived vectors and transformants

XX Disclosure; Page 4; 25pp; English.

XX Alpha amylase was purified from Pyrococcus furiosus and a partial
 CC amino acid sequence obtd. from a trypsin and cyanogen bromide
 CC digest. The sequence was used to design degenerate PCR primers
 CC which can be used to amplify the alpha amylase gene from the
 CC chromosomal DNA of P. furiosus. The recombinant alpha amylase

CC has excellent stability at high temperature so can be used for high
 CC temperature liquefaction of starch.
 CC See also AAR46558 and AAR47495.

SQ Sequence 12 AA;

Query Match 40.0%; Score 36; DB 15; Length 12;
 Best Local Similarity 85.7%; Pred. No. 8.2;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 IRQYVF 15
 :|||||
 Db 5 MRQYVF 11

RESULT 4

ID AAR53628 standard; protein; 12 AA.

XX AAR53628;
 AC AAR53628;

XX 07-JUL-1994 (first entry)

DE Pyrococcus furiosus alpha amylase fragment.

XX Pyrococcus furiosus; alpha amylase; liquefaction; polymers;
 KW glucopolymers; thermostable.
 XX Pyrococcus furiosus.

OS
 XX Key Location/Qualifiers
 FT Misc-difference 1
 FT /label= Gly or Met.

XX EP577257-A.

XX 05-JAN-1994.

XX 17-MAY-1993; 93EP-0303801.

XX 09-JUN-1992; 92US-0893928.

XX (UYJO) UNIV JOHNS HOPKINS.

XX Anfinsen CB, Laderman K;

XX WPI; 1994-009532/02.

XX Purified Pyrococcus furiosus alpha-amylase - used for the
 PT Industrial liquefaction of gluco-polymers at high temps.

XX Disclosure; Page 3; 41pp; English.

XX The purified Pyrococcus furiosus alpha amylase can act on substrates
 CC with a low degree of polymerisation. e.g. glucose polymers as short
 CC as maltotriose. The enzyme can be used for efficient industrial
 CC liquefaction of glucopolymers at high temperatures. This is given
 CC as a fragment of the alpha amylase although this sequence does not
 CC appear within the protein sequence of alpha amylase.

SQ Sequence 12 AA;

Query Match 40.0%; Score 36; DB 15; Length 12;
 Best Local Similarity 85.7%; Pred. No. 8.2;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 IRQYVF 15
 :|||||
 Db 5 MRQYVF 11

RESULT 5
 AAY6683

ID AAY6683 standard; peptide; 14 AA.

XX AAY6683;
 AC AAY6683;

XX 11-APR-2000 (first entry)

DE T cell antigen receptor Vbeta 8 chain peptide.

XX Rheumatoid arthritis; arthrosis deformans; T-cell antigen receptor;
 KW Vbeta chain; autoantigen; immunological tolerance.
 XX Homo sapiens.

XX WO9963084-A1.

XX 09-DEC-1999.

XX 28-MAY-1999; 99WO-JP02814.

XX 29-MAY-1998; 98JP-0149855.

XX 14-OCT-1998; 98JP-0328761.

XX (TORI) TORII PHARM CO LTD.

XX Nishioka K, Yoshino S;

XX WPI; 2000-086978/07.

XX N-PSDB; AA296613.

PT T-cell antigen receptor V-beta chain CDR3 region sequences accumulated
 in synovial membranes of rheumatoid arthritis patients -

XX Example 3; Page 86; 136pp; Japanese.

XX The invention relates to peptide sequences present in the synovial fluid
 CC and membranes of rheumatoid arthritis patients, arising from the CDR
 CC region of oligoclonal pathogenic T-cell antigen receptor Vbeta chains.
 CC Compositions which contain autoantigenic peptides binding specifically
 CC to T-cells expressing receptors containing the peptide sequences, which
 CC include antigen-specific immunological tolerance to rheumatoid arthritis
 CC can be used for the treatment and prevention of rheumatoid arthritis.
 CC The invention can be used for the diagnosis, treatment and prevention
 CC of rheumatoid arthritis. Sequences AAY66771-938 represent peptides from
 CC the various Vbeta chains of T cell antigen receptor.

SQ Sequence 14 AA;

Query Match 40.0%; Score 36; DB 21; Length 14;
 Best Local Similarity 54.5%; Pred. No. 9.9;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 ASTIRQYVFG 16
 ||:::| |||

Db 4 ASSLKDEQYFG 14

RESULT 6

AAR92123

ID AAR92123 standard; peptide; 25 AA.

XX AAR92123;

XX 15-AUG-1996 (first entry)

XX Streptomyces roseiscleroticus xylanase, xyl14, N-terminal peptide.

XX Wood pulp bleaching; xylanase; chromophores; toxic waste reduction;
 KW neutral pH; kraft pulp; paper.

XX Streptomyces roseiscleroticus NRRLB-11019.

XX US5498534-A.

PD 12-MAR-1996.
 XX
 PF 25-MAR-1992; 92US-0857060.
 XX
 PR 08-JUN-1994; 94US-0257965.
 PR 25-MAR-1992; 92US-0857060.
 PR 30-MAY-1995; 95US-0453289.
 XX
 PA (USDA) US SEC OF AGRIC.
 XX
 XX Elegir G, Grabski AC, Jeffries TW, Patel RN, Szakacs G;
 PI WPI; 1996-159688/16.
 XX
 DR Method for removing colour from kraft wood pulps - using xylanase
 PT isolated from Streptomyces roseiscleroticus
 PT
 XX Disclosure; Fig 6; 47pp; English.
 PS
 XX AAR92120-R92123 are N-terminal peptides of four xylanases, xyl1, xyl2,
 CC xyl3 and xyl4, encoded by Streptomyces roseiscleroticus strain
 CC NRRLB-11019. The xylanases may be used in a method for bleaching
 CC wood pulp which involves treating wood pulp with xylanase to
 CC release chromophores and extracting it to remove these chromophores.
 CC The wood pulp used is pref. kraft pulp from either soft or hard
 CC wood. Extraction of the chromophores may be performed using an
 CC alkali pref. hydrogen peroxide. Using xylanase as a bleaching
 CC agent reduces the need for using products such as chlorine which
 CC result in the formation of toxic degradation products which are
 CC difficult to remove by conventional waste treatment. Xylanase also
 CC works at a neutral pH, so a large pH shift of the pulp is not
 CC necessary.
 XX
 SQ Sequence 25 AA;
 Query Match 38.9%; Score 35; DB 17; Length 25;
 Best Local Similarity 47.1%; Pred. No. 31;
 Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 2 SKLEASTIRQYFYFGDA 18
 | | | : | | | |
 Db 3 STLGAARQSGYFCTA 19
 | | | : | | | |
 RESULT 7
 AAW80367
 ID AAW80367 standard; peptide; 25 AA.
 XX
 AC AAW80367;
 XX
 DT 13-JAN-1999 (first entry)
 XX
 DE N-terminal sequence of xylanase 4 (xyl4) of S. roseiscleroticus.
 XX
 KW Xylanase; colour removal; wood pulp; kraft pulp; softwood; hardwood.
 XX
 OS Streptomyces roseiscleroticus.
 XX
 PN US5834301-A.
 XX
 PD 10-NOV-1998.
 XX
 PF 18-DEC-1995; 95US-0574086.
 XX
 PR 08-JUN-1994; 94US-0257965.
 PR 25-MAR-1992; 92US-0857060.
 PR 30-MAY-1995; 95US-0453289.
 PR 18-DEC-1995; 95US-0574086.
 XX
 PA (USDA) US SEC OF AGRIC.
 XX
 XX Elegir G, Grabski AC, Jeffries TW, Patel RN, Szakacs G;
 PI WPI; 1996-159688/16.
 XX

DR WPI; 1999-008727/01.
 XX
 PT Removal of colour from wood pulp by treatment with specific
 PT xylanases - from Streptomyces strains NRRL 18982 or 18984 and then
 PT extracting the chromophores to reduce bleach demand in subsequent
 PT stages
 XX
 PS Example 1; Fig 6; 45pp; English.
 XX
 CC AAW80364-67 represent the N-terminal sequences of xylanases of
 CC S. roseiscleroticus strain NRRLB-11019. The xylanases are used in a
 CC method for removing colour from wood pulp. The method comprises
 CC treating with a xylanase, then extraction to remove chromophores.
 CC By removing chromophores (formed during the kraft cooking process),
 CC the method reduces the amount of bleach required in subsequent stages.
 CC The method is applied to kraft pulp from softwood or hardwood, or to
 CC pulp from secondary sources.
 XX
 SQ Sequence 25 AA;
 Query Match 38.9%; Score 35; DB 20; Length 25;
 Best Local Similarity 47.1%; Pred. No. 31;
 Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 2 SKLEASTIRQYFYFGDA 18
 | | | : | | | |
 Db 3 STLGAARQSGYFCTA 19
 | | | : | | | |
 RESULT 8
 AAR85526
 ID AAR85526 standard; peptide; 20 AA.
 XX
 AC AAR85526;
 XX
 DT 09-SEP-1996 (first entry)
 XX
 DE Varicella zoster gE glycoprotein residues 41-60.
 XX
 KW Glycoprotein; gE; VZV; antibody; gB; immunoreactive; immune response;
 KW infection; diagnosis; therapy.
 XX
 OS Varicella-zoster virus.
 XX
 PN WO9601900-A1.
 XX
 PD 25-JAN-1996.
 XX
 PF 03-JUL-1995; 95WO-GB01566.
 XX
 PR 07-JUL-1994; 94GB-0013751.
 XX
 PA (BRBI-) BRITISH BIOTECH PHARM LTD.
 XX
 PI Fowler WJ, Garcia-valcarcel Munoz-repiso M, Harper DR;
 PI Layton GT;
 XX
 DR WPI; 1996-097630/10.
 XX
 PT New isolated Varicella Zoster gE polypeptide(s) - used to develop
 PT products for use in vaccines, passive immunisation and diagnosis
 PT involving VZV infection
 XX
 PS Claim 3; Page 37; 47pp; English.
 XX
 CC AAR8522-R88549 represent fragments of the Varicella-Zoster virus (VZV)
 CC gE glycoprotein. This sequence represents residues 41-60 of gE. These
 CC sequences are used to create antibodies against the VZV gE glycoprotein.
 CC gE is one of six glycoproteins encoded by the VZV genome. From these
 CC six proteins, gE and gB are the major immunoreactive glycoproteins.
 CC These sequences can be used for stimulating an immune response against
 CC VZV infection. These peptides can also be used for determining the
 CC presence of anti-VZV gE antibodies in a sample, and in the diagnosis of

CC VZV infection. The antibodies against these sequences can be used for
 CC passive immunisation treatment, and in diagnostic applications. This
 CC sequence contains the major VZV gE immunodominant epitope and allows the
 CC development of products which can produce an enhanced and broader immune
 CC response.
 CC
 XX

SQ Sequence 20 AA;
 Query Match 37.8%; Score 34; DB 17; Length 20;
 Best Local Similarity 33.3%; Pred. NO. 35;
 Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 3 KLEASTIRQXYFGD 17
 ||: : : : ||| |
 Db 4 KLDNSVYEPYTHSD 18

RESULT 9
 AAW00599
 ID AAW00599 standard; peptide; 22 AA.
 XX
 AC AAW00599;
 XX
 DT 24-MAR-1997 (first entry)
 XX
 DE Heparan sulphate proteoglycan core protein fragment V18.
 XX
 KW Heparan sulphate proteoglycan; core protein; HSPG; V8 digestion; murine;
 KW tryptic peptide; neuroepithelial cell; c-myc; oncogene; embryonic cell;
 KW glycosaminoglycan; GAG; fibroblast growth factor; FGF; promoter; mouse;
 KW cytokine; cell proliferation; motor neuronal cell; neurological disease;
 KW FGF type IIIc receptor; Huntington's disease; Parkinson's disease;
 KW therapy.
 XX
 OS Mus musculus.
 XX
 PN WO9623003-A1.
 XX
 PD 01-AUG-1996.
 XX
 PF 25-JAN-1996; 96WO-AU00034.
 XX
 PR 16-JUN-1995; 95AU-0003560.
 PR 27-JAN-1995; 95AU-0000784.
 XX
 PA (AMRA-) AMRAD OPERATIONS PTY LTD.
 XX
 PI Bartlett PF, Nurcombe V;
 XX
 DR WPI; 1996-362636/36.
 XX
 PT Glycosaminoglycan polymer or fragment that interacts with cytokine,
 PT esp. FGF - used to rescue neuronal cells e.g. in transplants of
 PT FGF-expressing cells for treatment of neurological disease
 XX
 PS Example 1; Page 21; 57pp; English.
 XX
 CC AAW00599 and AAW00600 represent fragments of the heparan sulphate
 CC proteoglycan (HSPG) core protein. This sequence was generated from V8
 CC digestion of HSPG which was produced by the murine neuroepithelial cell
 CC line 2.3D. The 2.3D cell line is made by expressing the c-myc oncogene
 CC in cloned embryonic day 10 primary neuroepithelial cells. HSPG is used
 CC to prepare the glycosaminoglycan (GAG) polymers of the invention. The
 CC GAGs of the invention are able to interact with fibroblast growth
 CC factor-1 (FGF-1) or FGF-2, but not both. They are used to promote
 CC interaction between a specific cytokine and its target site on an animal
 CC cell. The GAGs can also be used to promote cell proliferation,
 CC migration, and/or differentiation of any tissue bearing the appropriate
 CC FGF receptor. The GAGs promote maintenance/survival of motor neuronal
 CC cells, and the viability of cells with the FGF type IIIc receptor. A
 CC particular application of the GAGs is to treat a neurological disease
 CC (especially Huntington's or Parkinson's), in conjunction with
 CC transplantation of cells that express FGF. Fragments of GAG that are

CC able to bind to FGF have antagonistic activity, so that they prevent the
 CC formation of an activating FGF-heparan sulphate-FGF receptor ternary
 CC complex. These HSPG core sequences are useful as substrates for
 CC synthesis of GAG, or for the production of antibodies, antagonists and
 CC agonists of GAG. Hybrids of GAG with a specific targeting peptide can
 CC be used for localised treatment.
 CC
 XX

SQ Sequence 22 AA;
 Query Match 36.7%; Score 33; DB 17; Length 22;
 Best Local Similarity 62.5%; Pred. NO. 60;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 11 QEYFEGDA 18
 | : : : | : |
 Db 8 QTFYFGDA 15

RESULT 10
 AAY66945
 ID AAY66945 standard; peptide; 20 AA.
 XX
 AC AAY66945;
 XX
 DT 11-APR-2000 (first entry)
 XX
 DE T cell antigen receptor Vbeta 14 chain peptide.
 XX
 KW Rheumatoid arthritis; arthrosis deformans; T-cell antigen-receptor;
 KW Vbeta chain; autoantigen; immunological tolerance.
 XX
 OS Homo sapiens.
 XX
 PN WO9963084-A1.
 XX
 PD 09-DEC-1999.
 XX
 PF 28-MAY-1999; 99WO-JP02814.
 XX
 PR 29-MAY-1998; 98JP-0149855.
 PR 14-OCT-1998; 98JP-0328761.
 XX
 PA (TORI) TORII PHARM CO LTD.
 XX
 PI Nishioka K, Yoshino S;
 XX
 DR WPI; 2000-086978/07.
 DR N-PSDB; AA96675.
 XX
 PT T-cell antigen receptor V-beta chain CDR3 region sequences accumulated
 PT in synovial membranes of rheumatoid arthritis patients -
 XX
 PS Example 3; Page 124; 136pp; Japanese.
 XX
 CC The invention relates to peptide sequences present in the synovial fluid
 CC and membranes of rheumatoid arthritis patients, arising from the CDR
 CC region of oligoclonal pathogenic T-cell antigen receptor Vbeta chains.
 CC Compositions which contain autoantigenic peptides binding specifically
 CC to T-cells expressing receptors containing the peptide sequences, which
 CC include antigen-specific immunological tolerance to rheumatoid arthritis.
 CC can be used for the treatment and prevention of rheumatoid arthritis.
 CC The invention can be used for the diagnosis, treatment and prevention
 CC of rheumatoid arthritis. Sequences AAY66771-958 represent peptides from
 CC the various Vbeta chains of T cell antigen receptor.

XX
 SQ Sequence 20 AA;
 Query Match 35.6%; Score 32; DB 21; Length 20;
 Best Local Similarity 46.7%; Pred. NO. 81;
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 2 SKLEASTIRQXYFG 16
 | : : | : | : |

Db 5 SSLKAGGHRSEQFFG 19

RESULT 11
ABG62324
ID ABG62324 standard; Peptide; 25 AA.
XX
XX
AC ABG62324;
XX
XX
DT 21-AUG-2002 (first entry)
XX
XX
DE Eubacterial DNA polymerase IV QLSLF motif containing peptide #54.
XX
XX
KW DNA polymerase III; beta subunit; eubacteria; antibacterial;
KW eubacterial infection.
XX
XX
OS Mycobacterium tuberculosis.
XX
XX
PN WO200238596-A1.
XX
XX
PD 16-MAY-2002.
XX
XX
PF 08-NOV-2001; 2001WO-AU01436.
XX
XX
PR 08-NOV-2000; 2000AU-0001320.
PR 06-FEB-2001; 2001AU-0002919.
XX
XX
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
XX
XX
PI Dalrymple BP, Kongsuwan K, Wijffels GL, Jennings PA, Kemp GW;
XX
XX
DR WPI; 2002-471546/50.
XX
XX
PT New molecule having surface analogous to surface of domain of
PT eubacterial beta protein contacted by proteins that interact with beta
PT protein, useful to identify inhibitors of beta protein-ligand
PT interaction
XX
XX
PS Example 1; Page 27; 326pp; English.

The invention relates to a molecule (I) comprising a surface (S)
analogous to the surface of the domain of eubacterial beta protein
contacted by proteins that interact with beta protein, where the
surface is defined by the residues X(170), X(172), X(175), X(177),
X(241), X(242), X(247), X(346), and X(362), where the
superscript numbers designate the position of residues in Escherichia
coli beta protein, or the equivalent residues in homologues from other
species of eubacteria, and where:
X(170) = Val, Ile, Ala, Thr, Ser or Glu; X(172) = Thr, Ser or Ile;
X(175) = His, Tyr, Phe, Lys, Ile, Gln or Arg; X(177) = Leu, Met, Ile,
Phe, Val or Ala; X(241) = Phe, Tyr or Leu; X(242) = Pro, Leu or Ile;
X(247) = Val, Ile, Ala, Phe, Leu or Met; X(346) = Ser, Pro, Ala, Tyr or
Lys; X(360) = Ile, Leu or Val; and X(362) = Met, Leu, Val, Ser, Thr or
Arg. Also included are methods of identifying a modulator of the
interaction between a eubacterial beta protein and proteins that interact
with them, reducing (M4) the effect of eubacterial infestation of a
biological system, involves delivering to a system infested with a
eubacterial species, a modulator of the interaction between eubacterial
beta protein and proteins that interact with the beta protein; and
(4) a template (II) for the design of a compound that binds to at least
part of (S) of beta protein as defined above comprises a (P) such as
X¹-IX², X³-IX², X³-IX²-IX⁴, GlnX⁵-X³-IX², GlnX⁵-X³-IX², GlnX⁵-X³-IX²,
where: x = any amino acid residue; X¹ = Leu, Met, Ile, or Phe;
X² = Leu, Ile, Val, Cys, Phe, Tyr, Trp, Pro, Asp, Ala or Gly;
X³ = Ala, Gly, Thr, Asp, Ser, or Pro; X⁴ = Ala or Gly; X⁵ = Leu;
and X⁶ = Leu, Ile, Val, Cys, Phe, Tyr, Trp or Pro. The method are useful
for identifying a modulator of the interaction between a eubacterial beta
protein and proteins that interact with the beta protein. (M4) is useful
for reducing the effect of eubacterial infestation of a biological
system. The compounds identified using above mentioned methods are
useful as antibacterial agent for treatment or prevention of disease in
humans, animals and plants. The present sequence is a eubacterial
peptide from a DNA binding protein or polymerase which contains a DNA

CC polymerase III beta subunit binding site.
XX
SQ Sequence 25 AA;
Query Match 35.6%; Score 32; DB 23; Length 25;
Best Local Similarity 58.3%; Pred. No. 1.le-02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
OY 7 STIROEYVFGDA 18
| | | | | | | | | |
Db 7 SDIROESLPADS 18
| | | | | | | | | |
RESULT 12
ABP47562
ID ABP47562 standard; Peptide; 9 AA.
XX
XX
AC ABP47562;
XX
XX
DT 19-AUG-2002 (first entry)
XX
XX
DE N. meningitidis LOS peptidic mimotope related peptide SEQ ID NO:138.
XX
XX
KW Neisseria meningitidis; meningococcus; meningococcal; vaccine; LOS;
KW lipooligosaccharide; monoclonal antibody; antibacterial; infection;
KW antiinflammatory; meningococcal disease.
XX
XX
OS Neisseria meningitidis.
OS Synthetic.
XX
XX
PN WO200228888-A2.
XX
XX
PD 11-APR-2002.
XX
XX
PF 03-OCT-2001; 2001WO-EP11409.
XX
XX
PR 03-OCT-2000; 2000GB-0024200.
XX
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX
PI De Bolle XT, Letesson J, Lobet Y, Mertens PY, Poolman J, Voet P;
XX
XX
DR WPI; 2002-479596/51.
XX
XX
PT Novel mimotope of Neisseria meningitidis surface, for treating
PT meningococcal disease, comprising a peptide epitope obtainable by
PT screening peptide library with a specific monoclonal antibody
XX
XX
PS Claim 3; Page 42; 53pp; English.
XX
XX
CC The present invention describes mimotopes (I) of a surface L3, 7, 9,
CC of meningococcal lipopolysaccharides (LOS) of Neisseria meningitidis
CC comprising a peptide epitope obtained by screening a peptide library
CC with a monoclonal antibody (MAB) like 4B12C10, H44/24, H44/58, H44/70
CC or H44/78. (I) is antigenically cross-reactive with MAB. (I) have
CC antibacterial and antiinflammatory activities, and can be used in
CC vaccines. MAB is useful in the identification of (I). (I) or MAB are
CC useful as a medicament, and also in the manufacture of a medicament for
CC treating or preventing meningococcal disease. (I) and MAB are useful
CC for treating a patient suffering from or susceptible to meningococcal
CC disease by administering (I) or MAB to the patient. (I) is useful in
CC a diagnostic assay for meningococcal infection to detect antibodies
CC against L3, 7, 9, LOS and to detect the presence of L3, 7, 9 immunotype
CC meningococcus in a sample from a patient. ABN88464 to ABN88487 and
CC ABP47336 to ABP47754 represent sequences used in the exemplification
CC of the present invention.
XX
XX
SQ Sequence 9 AA;
Query Match 34.4%; Score 31; DB 23; Length 9;
Best Local Similarity 55.6%; Pred. No. 7.8e-05;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 6 ASTIRQY 14
I : I : I :
Db 1 AKTVRGDYY 9

RESULT 13

ABP47702
ID ABP47702 standard; Peptide; 11 AA.

XX AC
XX ABP47702;

DT 19-AUG-2002 (first entry)

XX N. meningitidis LOS peptidic mimotope related peptide SEQ ID NO:278.
XX Neisseria meningitidis; meningococcus; meningococcal; vaccine; LOS;
KW lipooligosaccharide; monoclonal antibody; antibacterial; infection;
KW antinflammatory; meningococcal disease.

XX OS
XX Neisseria meningitidis.
XX Synthetic.

XX WO200228888-A2.

XX 11-APR-2002.

XX 03-OCT-2001; 2001WO-EP111409.

XX 03-OCT-2000; 2000GB-0024200.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX De Bolle XT, Letesson J, Lobet Y, Mertens PY, Poolman J, Voet P;
XX WPI; 2002-479596/51.

XX Novel mimotope of Neisseria meningitidis surface, for treating
PT meningococcal disease, comprising a peptide epitope obtainable by
PT screening peptide library with a specific monoclonal antibody
XX Claim 15; Page 45; 55pp; English.

XX The present invention describes mimotopes (I) of a surface L3, 7, 9,
CC of meningococcal lipopolysaccharides (LOS) of Neisseria meningitidis
CC comprising a peptide epitope obtained by screening a peptide library
CC with a monoclonal antibody (MAB) like 4B12C10, H44/24, H44/58, H44/70
CC or H44/78. (I) is antigenically cross-reactive with MAB. (I) have
CC antibacterial and antiinflammatory activities, and can be used in
CC vaccines. MAB is useful in the identification of (I). (I) or MAB are
CC useful as a medicament, and also in the manufacture of a medicament for
CC treating or preventing meningococcal disease. (I) and MAB are useful
CC for treating a patient suffering from or susceptible to meningococcal
CC disease by administering (I) or MAB to the patient. (I) is useful in
CC a diagnostic assay for meningococcal infection to detect antibodies
CC against L3, 7, 9, LOS and to detect the presence of L3, 7, 9 immunotype
CC meningococcus in a sample from a patient. ABN88464 to ABN88487 and
CC ABP47336 to ABP47754 represent sequences used in the exemplification
CC of the present invention.

XX Sequence 11 AA;

Query Match 34.4%; Score 31; DB 23; Length 11;
Best Local Similarity 55.6%; Pred. No. 59;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 6 ASTIRQY 14
I : I : I :
Db 2 AKTVRGDYY 10

RESULT 14

ABG62293
ID ABG62293 standard; Peptide; 19 AA.

XX AC

XX ABG62293;

XX 21-AUG-2002 (first entry)

XX Eubacterial DNA polymerase IV QLSLF motif containing peptide #23.

XX DNA polymerase III; beta subunit; eubacteria; antibacterial;
KW eubacterial infection.

XX Bordetella pertussis.

XX WO200238596-A1.

XX 16-MAY-2002.

XX 08-NOV-2001; 2001WO-AU01436.

XX 08-NOV-2000; 2000AU-0001320.

XX 06-FEB-2001; 2001AU-0002919.

XX (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX Dalrymple BP, Kongsuwan K, Wijffels GL, Jennings PA, Kemp GW;

XX WPI; 2002-471546/50.

XX New molecule having surface analogous to surface of domain of
PT eubacterial beta protein contacted by proteins that interact with beta
PT protein, useful to identify inhibitors of beta protein-ligand
PT interaction

XX Example 1; Page 26; 326pp; English.

XX The invention relates to a molecule (I) comprising a surface (S)
CC analogous to the surface of the domain of eubacterial beta protein
CC contacted by proteins that interact with beta protein, where the
CC surface is defined by the residues X(170), X(172), X(175), X(177),
CC X(241), X(242), X(247), X(346), X(360), and X(362), where the
CC superscript numbers designate the position of residues in Escherichia
CC coli beta protein, or the equivalent residues in homologues from other
CC species of eubacteria, and where:
CC X(170) = Val, Ile, Ala, Thr, Ser or Glu; X(172) = Thr, Ser or Ile;
CC X(175) = His, Tyr, Phe, Lys, Ile, Gln or Arg; X(177) = Leu, Met, Ile,
CC Phe, Val or Ala; X(241) = Phe, Tyr or Leu; X(242) = Pro, Leu or Ile;
CC X(247) = Val, Ile, Ala, Phe, Leu or Met; X(346) = Ser, Pro, Ala, Tyr or
CC Lys; X(360) = Ile, Leu or Val; and X(362) = Met, Leu, Val, Ser, Thr or
CC Arg. Also included are methods of identifying a modulator of the
CC interaction between a eubacterial beta protein and proteins that interact
CC with them, reducing (M4) the effect of eubacterial infestation of a
CC biological system, involves delivering to a system infested with a
CC eubacterial species, a modulator of the interaction between eubacterial
CC beta protein and proteins that interact with the beta protein; and
CC (4) a template (II) for the design of a compound that binds to at least
CC part of (S) of beta protein as defined above comprises a (P) such as
CC X¹X²X³X⁴X⁵X⁶X⁷X⁸X⁹X¹⁰X¹¹X¹²X¹³X¹⁴X¹⁵X¹⁶X¹⁷X¹⁸X¹⁹X²⁰X²¹X²²X²³X²⁴X²⁵X²⁶X²⁷X²⁸X²⁹X³⁰X³¹X³²X³³X³⁴X³⁵X³⁶X³⁷X³⁸X³⁹X⁴⁰X⁴¹X⁴²X⁴³X⁴⁴X⁴⁵X⁴⁶X⁴⁷X⁴⁸X⁴⁹X⁵⁰X⁵¹X⁵²X⁵³X⁵⁴X⁵⁵X⁵⁶X⁵⁷X⁵⁸X⁵⁹X⁶⁰X⁶¹X⁶²X⁶³X⁶⁴X⁶⁵X⁶⁶X⁶⁷X⁶⁸X⁶⁹X⁷⁰X⁷¹X⁷²X⁷³X⁷⁴X⁷⁵X⁷⁶X⁷⁷X⁷⁸X⁷⁹X⁸⁰X⁸¹X⁸²X⁸³X⁸⁴X⁸⁵X⁸⁶X⁸⁷X⁸⁸X⁸⁹X⁹⁰X⁹¹X⁹²X⁹³X⁹⁴X⁹⁵X⁹⁶X⁹⁷X⁹⁸X⁹⁹X¹⁰⁰X¹⁰¹X¹⁰²X¹⁰³X¹⁰⁴X¹⁰⁵X¹⁰⁶X¹⁰⁷X¹⁰⁸X¹⁰⁹X¹¹⁰X¹¹¹X¹¹²X¹¹³X¹¹⁴X¹¹⁵X¹¹⁶X¹¹⁷X¹¹⁸X¹¹⁹X¹²⁰X¹²¹X¹²²X¹²³X¹²⁴X¹²⁵X¹²⁶X¹²⁷X¹²⁸X¹²⁹X¹³⁰X¹³¹X¹³²X¹³³X¹³⁴X¹³⁵X¹³⁶X¹³⁷X¹³⁸X¹³⁹X¹⁴⁰X¹⁴¹X¹⁴²X¹⁴³X¹⁴⁴X¹⁴⁵X¹⁴⁶X¹⁴⁷X¹⁴⁸X¹⁴⁹X¹⁵⁰X¹⁵¹X¹⁵²X¹⁵³X¹⁵⁴X¹⁵⁵X¹⁵⁶X¹⁵⁷X¹⁵⁸X¹⁵⁹X¹⁶⁰X¹⁶¹X¹⁶²X¹⁶³X¹⁶⁴X¹⁶⁵X¹⁶⁶X¹⁶⁷X¹⁶⁸X¹⁶⁹X¹⁷⁰X¹⁷¹X¹⁷²X¹⁷³X¹⁷⁴X¹⁷⁵X¹⁷⁶X¹⁷⁷X¹⁷⁸X¹⁷⁹X¹⁸⁰X¹⁸¹X¹⁸²X¹⁸³X¹⁸⁴X¹⁸⁵X¹⁸⁶X¹⁸⁷X¹⁸⁸X¹⁸⁹X¹⁹⁰X¹⁹¹X¹⁹²X¹⁹³X¹⁹⁴X¹⁹⁵X¹⁹⁶X¹⁹⁷X¹⁹⁸X¹⁹⁹X²⁰⁰X²⁰¹X²⁰²X²⁰³X²⁰⁴X²⁰⁵X²⁰⁶X²⁰⁷X²⁰⁸X²⁰⁹X²¹⁰X²¹¹X²¹²X²¹³X²¹⁴X²¹⁵X²¹⁶X²¹⁷X²¹⁸X²¹⁹X²²⁰X²²¹X²²²X²²³X²²⁴X²²⁵X²²⁶X²²⁷X²²⁸X²²⁹X²³⁰X²³¹X²³²X²³³X²³⁴X²³⁵X²³⁶X²³⁷X²³⁸X²³⁹X²⁴⁰X²⁴¹X²⁴²X²⁴³X²⁴⁴X²⁴⁵X²⁴⁶X²⁴⁷X²⁴⁸X²⁴⁹X²⁵⁰X²⁵¹X²⁵²X²⁵³X²⁵⁴X²⁵⁵X²⁵⁶X²⁵⁷X²⁵⁸X²⁵⁹X²⁶⁰X²⁶¹X²⁶²X²⁶³X²⁶⁴X²⁶⁵X²⁶⁶X²⁶⁷X²⁶⁸X²⁶⁹X²⁷⁰X²⁷¹X²⁷²X²⁷³X²⁷⁴X²⁷⁵X²⁷⁶X²⁷⁷X²⁷⁸X²⁷⁹X²⁸⁰X²⁸¹X²⁸²X²⁸³X²⁸⁴X²⁸⁵X²⁸⁶X²⁸⁷X²⁸⁸X²⁸⁹X²⁹⁰X²⁹¹X²⁹²X²⁹³X²⁹⁴X²⁹⁵X²⁹⁶X²⁹⁷X²⁹⁸X²⁹⁹X³⁰⁰X³⁰¹X³⁰²X³⁰³X³⁰⁴X³⁰⁵X³⁰⁶X³⁰⁷X³⁰⁸X³⁰⁹X³¹⁰X³¹¹X³¹²X³¹³X³¹⁴X³¹⁵X³¹⁶X³¹⁷X³¹⁸X³¹⁹X³²⁰X³²¹X³²²X³²³X³²⁴X³²⁵X³²⁶X³²⁷X³²⁸X³²⁹X³³⁰X³³¹X³³²X³³³X³³⁴X³³⁵X³³⁶X³³⁷X³³⁸X³³⁹X³⁴⁰X³⁴¹X³⁴²X³⁴³X³⁴⁴X³⁴⁵X³⁴⁶X³⁴⁷X³⁴⁸X³⁴⁹X³⁵⁰X³⁵¹X³⁵²X³⁵³X³⁵⁴X³⁵⁵X³⁵⁶X³⁵⁷X³⁵⁸X³⁵⁹X³⁶⁰X³⁶¹X³⁶²X³⁶³X³⁶⁴X³⁶⁵X³⁶⁶X³⁶⁷X³⁶⁸X³⁶⁹X³⁷⁰X³⁷¹X³⁷²X³⁷³X³⁷⁴X³⁷⁵X³⁷⁶X³⁷⁷X³⁷⁸X³⁷⁹X³⁸⁰X³⁸¹X³⁸²X³⁸³X³⁸⁴X³⁸⁵X³⁸⁶X³⁸⁷X³⁸⁸X³⁸⁹X³⁹⁰X³⁹¹X³⁹²X³⁹³X³⁹⁴X³⁹⁵X³⁹⁶X³⁹⁷X³⁹⁸X³⁹⁹X⁴⁰⁰X⁴⁰¹X⁴⁰²X⁴⁰³X⁴⁰⁴X⁴⁰⁵X⁴⁰⁶X⁴⁰⁷X⁴⁰⁸X⁴⁰⁹X⁴¹⁰X⁴¹¹X⁴¹²X⁴¹³X⁴¹⁴X⁴¹⁵X⁴¹⁶X⁴¹⁷X⁴¹⁸X⁴¹⁹X⁴²⁰X⁴²¹X⁴²²X⁴²³X⁴²⁴X⁴²⁵X⁴²⁶X⁴²⁷X⁴²⁸X⁴²⁹X⁴³⁰X⁴³¹X⁴³²X⁴³³X⁴³⁴X⁴³⁵X⁴³⁶X⁴³⁷X⁴³⁸X⁴³⁹X⁴⁴⁰X⁴⁴¹X⁴⁴²X⁴⁴³X⁴⁴⁴X⁴⁴⁵X⁴⁴⁶X⁴⁴⁷X⁴⁴⁸X⁴⁴⁹X⁴⁵⁰X⁴⁵¹X⁴⁵²X⁴⁵³X⁴⁵⁴X⁴⁵⁵X⁴⁵⁶X⁴⁵⁷X⁴⁵⁸X⁴⁵⁹X⁴⁶⁰X⁴⁶¹X⁴⁶²X⁴⁶³X⁴⁶⁴X⁴⁶⁵X⁴⁶⁶X⁴⁶⁷X⁴⁶⁸X⁴⁶⁹X⁴⁷⁰X⁴⁷¹X⁴⁷²X⁴⁷³X⁴⁷⁴X⁴⁷⁵X⁴⁷⁶X⁴⁷⁷X⁴⁷⁸X⁴⁷⁹X⁴⁸⁰X⁴⁸¹X⁴⁸²X⁴⁸³X⁴⁸⁴X⁴⁸⁵X⁴⁸⁶X⁴⁸⁷X⁴⁸⁸X⁴⁸⁹X⁴⁹⁰X⁴⁹¹X⁴⁹²X⁴⁹³X⁴⁹⁴X⁴⁹⁵X⁴⁹⁶X⁴⁹⁷X⁴⁹⁸X⁴⁹⁹X⁵⁰⁰X⁵⁰¹X⁵⁰²X⁵⁰³X⁵⁰⁴X⁵⁰⁵X⁵⁰⁶X⁵⁰⁷X⁵⁰⁸X⁵⁰⁹X⁵¹⁰X⁵¹¹X⁵¹²X⁵¹³X⁵¹⁴X⁵¹⁵X⁵¹⁶X⁵¹⁷X⁵¹⁸X⁵¹⁹X⁵²⁰X⁵²¹X⁵²²X⁵²³X⁵²⁴X⁵²⁵X⁵²⁶X⁵²⁷X⁵²⁸X⁵²⁹X⁵³⁰X⁵³¹X⁵³²X⁵³³X⁵³⁴X⁵³⁵X⁵³⁶X⁵³⁷X⁵³⁸X⁵³⁹X⁵⁴⁰X⁵⁴¹X⁵⁴²X⁵⁴³X⁵⁴⁴X⁵⁴⁵X⁵⁴⁶X⁵⁴⁷X⁵⁴⁸X⁵⁴⁹X⁵⁵⁰X⁵⁵¹X⁵⁵²X⁵⁵³X⁵⁵⁴X⁵⁵⁵X⁵⁵⁶X⁵⁵⁷X⁵⁵⁸X⁵⁵⁹X⁵⁶⁰X⁵⁶¹X⁵⁶²X⁵⁶³X⁵⁶⁴X⁵⁶⁵X⁵⁶⁶X⁵⁶⁷X⁵⁶⁸X⁵⁶⁹X⁵⁷⁰X⁵⁷¹X⁵⁷²X⁵⁷³X⁵⁷⁴X⁵⁷⁵X⁵⁷⁶X⁵⁷⁷X⁵⁷⁸X⁵⁷⁹X⁵⁸⁰X⁵⁸¹X⁵⁸²X⁵⁸³X⁵⁸⁴X⁵⁸⁵X⁵⁸⁶X⁵⁸⁷X⁵⁸⁸X⁵⁸⁹X⁵⁹⁰X⁵⁹¹X⁵⁹²X⁵⁹³X⁵⁹⁴X⁵⁹⁵X⁵⁹⁶X⁵⁹⁷X⁵⁹⁸X⁵⁹⁹X⁶⁰⁰X⁶⁰¹X⁶⁰²X⁶⁰³X⁶⁰⁴X⁶⁰⁵X⁶⁰⁶X⁶⁰⁷X⁶⁰⁸X⁶⁰⁹X⁶¹⁰X⁶¹¹X⁶¹²X⁶¹³X⁶¹⁴X⁶¹⁵X⁶¹⁶X⁶¹⁷X⁶¹⁸X⁶¹⁹X⁶²⁰X⁶²¹X⁶²²X⁶²³X⁶²⁴X⁶²⁵X⁶²⁶X⁶²⁷X⁶²⁸X⁶²⁹X⁶³⁰X⁶³¹X⁶³²X⁶³³X⁶³⁴X⁶³⁵X⁶³⁶X⁶³⁷X⁶³⁸X⁶³⁹X⁶⁴⁰X⁶⁴¹X⁶⁴²X⁶⁴³X⁶⁴⁴X⁶⁴⁵X⁶⁴⁶X⁶⁴⁷X⁶⁴⁸X⁶⁴⁹X⁶⁵⁰X⁶⁵¹X⁶⁵²X⁶⁵³X⁶⁵⁴X⁶⁵⁵X⁶⁵⁶X⁶⁵⁷X⁶⁵⁸X⁶⁵⁹X⁶⁶⁰X⁶⁶¹X⁶⁶²X⁶⁶³X⁶⁶⁴X⁶⁶⁵X⁶⁶⁶X⁶⁶⁷X⁶⁶⁸X⁶⁶⁹X⁶⁷⁰X⁶⁷¹X⁶⁷²X⁶⁷³X⁶⁷⁴X⁶⁷⁵X⁶⁷⁶X⁶⁷⁷X⁶⁷⁸X⁶⁷⁹X⁶⁸⁰X⁶⁸¹X⁶⁸²X⁶⁸³X⁶⁸⁴X⁶⁸⁵X⁶⁸⁶X⁶⁸⁷X⁶⁸⁸X⁶⁸⁹X⁶⁹⁰X⁶⁹¹X⁶⁹²X⁶⁹³X⁶⁹⁴X⁶⁹⁵X⁶⁹⁶X⁶⁹⁷X⁶⁹⁸X⁶⁹⁹X⁷⁰⁰X⁷⁰¹X⁷⁰²X⁷⁰³X⁷⁰⁴X⁷⁰⁵X⁷⁰⁶X⁷⁰⁷X⁷⁰⁸X⁷⁰⁹X⁷¹⁰X⁷¹¹X⁷¹²X⁷¹³X⁷¹⁴X⁷¹⁵X⁷¹⁶X⁷¹⁷X⁷¹⁸X⁷¹⁹X⁷²⁰X⁷²¹X⁷²²X⁷²³X⁷²⁴X⁷²⁵X⁷²⁶X⁷²⁷X⁷²⁸X⁷²⁹X⁷³⁰X⁷³¹X⁷³²X⁷³³X⁷³⁴X⁷³⁵X⁷³⁶X⁷³⁷X⁷³⁸X⁷³⁹X⁷⁴⁰X⁷⁴¹X⁷⁴²X⁷⁴³X⁷⁴⁴X⁷⁴⁵X⁷⁴⁶X⁷⁴⁷X⁷⁴⁸X⁷⁴⁹X⁷⁵⁰X⁷⁵¹X⁷⁵²X⁷⁵³X⁷⁵⁴X⁷⁵⁵X⁷⁵⁶X⁷⁵⁷X⁷⁵⁸X⁷⁵⁹X⁷⁶⁰X⁷⁶¹X⁷⁶²X⁷⁶³X⁷⁶⁴X⁷⁶⁵X⁷⁶⁶X⁷⁶⁷X⁷⁶⁸X⁷⁶⁹X⁷⁷⁰X⁷⁷¹X⁷⁷²X⁷⁷³X⁷⁷⁴X⁷⁷⁵X⁷⁷⁶X⁷⁷⁷X⁷⁷⁸X⁷⁷⁹X⁷⁸⁰X⁷⁸¹X⁷⁸²X⁷⁸³X⁷⁸⁴X⁷⁸⁵X⁷⁸⁶X⁷⁸⁷X⁷⁸⁸X⁷⁸⁹X⁷⁹⁰X⁷⁹¹X⁷⁹²X⁷⁹³X⁷⁹⁴X⁷⁹⁵X⁷⁹⁶X⁷⁹⁷X⁷⁹⁸X⁷⁹⁹X⁸⁰⁰X⁸⁰¹X⁸⁰²X⁸⁰³X⁸⁰⁴X⁸⁰⁵X⁸⁰⁶X⁸⁰⁷X⁸⁰⁸X⁸⁰⁹X⁸¹⁰X⁸¹¹X⁸¹²X⁸¹³X⁸¹⁴X⁸¹⁵X⁸¹⁶X⁸¹⁷X⁸¹⁸X⁸¹⁹X⁸²⁰X⁸²¹X⁸²²X⁸²³X⁸²⁴X⁸²⁵X⁸²⁶X⁸²⁷X⁸²⁸X⁸²⁹X⁸³⁰X⁸³¹X⁸³²X⁸³³X⁸³⁴X⁸³⁵X⁸³⁶X⁸³⁷X⁸³⁸X⁸³⁹X⁸⁴⁰X⁸⁴¹X⁸⁴²X⁸⁴³X⁸⁴⁴X⁸⁴⁵X⁸⁴⁶X⁸⁴⁷X⁸⁴⁸X⁸⁴⁹X⁸⁵⁰X⁸⁵¹X⁸⁵²X⁸⁵³X⁸⁵⁴X⁸⁵⁵X⁸⁵⁶X⁸⁵⁷X⁸⁵⁸X⁸⁵⁹X⁸⁶⁰X⁸⁶¹X⁸⁶²X⁸⁶³X⁸⁶⁴X⁸⁶⁵X⁸⁶⁶X⁸⁶⁷X⁸⁶⁸X⁸⁶⁹X⁸⁷⁰X⁸⁷¹X⁸⁷²X⁸⁷³X⁸⁷⁴X⁸⁷⁵X⁸⁷⁶X⁸⁷⁷X⁸⁷⁸X⁸⁷⁹X⁸⁸⁰X⁸⁸¹X⁸⁸²X⁸⁸³X⁸⁸⁴X⁸⁸⁵X⁸⁸⁶X⁸⁸⁷X⁸⁸⁸X⁸⁸⁹X⁸⁹⁰X⁸⁹¹X⁸⁹²X⁸⁹³X⁸⁹⁴X⁸⁹⁵X⁸⁹⁶X⁸⁹⁷X⁸⁹⁸X⁸⁹⁹X⁹⁰⁰X⁹⁰¹X⁹⁰²X⁹⁰³X⁹⁰⁴X⁹⁰⁵X⁹⁰⁶X⁹⁰⁷X⁹⁰⁸X⁹⁰⁹X⁹¹⁰X⁹¹¹X⁹¹²X⁹¹³X⁹¹⁴X⁹¹⁵X⁹¹⁶X⁹¹⁷X⁹¹⁸X⁹¹⁹X⁹²⁰X⁹²¹X⁹²²X⁹²³X⁹²⁴X⁹²⁵X⁹²⁶X⁹²⁷X⁹²⁸X⁹²⁹X⁹³⁰X⁹³¹X⁹³²X⁹³³X⁹³⁴X⁹³⁵X⁹³⁶X⁹³⁷X⁹³⁸X⁹³⁹X⁹⁴⁰X⁹⁴¹X⁹⁴²X⁹⁴³X⁹⁴⁴X⁹⁴⁵X⁹⁴⁶X⁹⁴⁷X⁹⁴⁸X⁹⁴⁹X⁹⁵⁰X⁹⁵¹X⁹⁵²X⁹⁵³X⁹⁵⁴X⁹⁵⁵X⁹⁵⁶X⁹⁵⁷X⁹⁵⁸X⁹⁵⁹X⁹⁶⁰X⁹⁶¹X⁹⁶²X⁹⁶³X⁹⁶⁴X⁹⁶⁵X⁹⁶⁶X⁹⁶⁷X⁹⁶⁸X⁹⁶⁹X⁹⁷⁰X⁹⁷¹X⁹⁷²X⁹⁷³X⁹⁷⁴X⁹⁷⁵X⁹⁷⁶X⁹⁷⁷X⁹⁷⁸X⁹⁷⁹X⁹⁸⁰X⁹⁸¹X⁹⁸²X⁹⁸³X⁹⁸⁴X⁹⁸⁵X⁹⁸⁶X⁹⁸⁷X⁹⁸⁸X⁹⁸⁹X⁹⁹⁰X⁹⁹¹X⁹⁹²X⁹⁹³X⁹⁹⁴X⁹⁹⁵X⁹⁹⁶X⁹⁹⁷X⁹⁹⁸X⁹⁹⁹X¹⁰⁰⁰X¹⁰⁰¹X¹⁰⁰²X¹⁰⁰³X¹⁰⁰⁴X¹⁰⁰⁵X¹⁰⁰⁶X¹⁰⁰⁷X¹⁰⁰⁸X¹⁰⁰⁹X¹⁰¹⁰X¹⁰¹¹X¹⁰¹²X¹⁰¹³X¹⁰¹⁴X¹⁰¹⁵X¹⁰¹⁶X¹⁰¹⁷X¹⁰¹⁸X¹⁰¹⁹X¹⁰²⁰X¹⁰²¹X¹⁰²²X¹⁰²³X¹⁰²⁴X¹⁰²⁵X¹⁰²⁶X¹⁰²⁷X¹⁰²⁸X¹⁰²⁹X¹⁰³⁰X

Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 EASTIRQYYFGDA 18
: | | | |
Db 5 QAEAPROAELEFGDA 18

RESULT 15

AAP71322
ID AAP71322 standard; peptide; 20 AA.

XX

AC AAP71322;

XX

DT 07-MAY-1991 (first entry)

XX

DE Phosphopeptide 3.

XX

KW Caries; gingivitis; periodontal disease; osteoporosis; osteomalacia.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Modified-site 7

FT FT /label= phosphoserine

FT FT 8

FT FT /label= phosphoserine

FT FT 9

FT FT /label= phosphoserine

FT FT 15

FT FT /label= phosphoserine

XX

PN WO8707616-A.

XX

XX

PD 17-DEC-1987.

XX

PF 12-JUN-1987; 87WO-AU00172.

XX

PR 12-JUN-1986; 86AU-0006385.

XX

PA (VICT-) VICTORIA DAIRY INDUSTRY AUTHORITY.

PA (UYME-) UNIVERSITY OF MELBOURNE.

XX

PA (REYN/) EC REYNOLDS.

XX

PI Reynolds EC;

XX

DR WPI; 1987-362707/51.

XX

PT New phosphopeptides contg. defined amino acid sequence - useful in treatment of dental, rarefying bone diseases and disease relating to malabsorption of minerals.

XX

PS Claim 6; Page 17; 22pp; English.

XX

CC The phosphopeptide is used in compans. at a conc. of 0.01-5 wt%.

XX

CC See also AAP71320-P71324.

XX

SQ Sequence 20 AA;

XX

Query Match 34.4%; Score 31; DB 8; Length 20;

Best Local Similarity 57.1%; Pred. No. 1.2e+02;

Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 VSKLEASTIRQYY 14

XX

XX

Db 6 VSSSEESIISQETY 19

Search completed: April 23, 2003, 13:43:02

Job time : 29.2022 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:21:27 ; Search time 27.9101 Seconds
(without alignments)
132.885 Million cell updates/sec

Title: US-09-836-073-17

Perfect score: 95

Sequence: 1 DDADQRIKQLEYFGNI 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_morganelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	96.8	396	5 O01806	O01806 caenorhabdi
2	57	60.0	411	10 Q9FL36	Q9FL36 arabidopsi
3	57	60.0	422	10 Q9A38	Q9A38 arabidopsi
4	54	56.8	545	10 O80367	O80367 arabidopsi
5	53	55.8	389	10 Q8S0T8	Q8S0T8 oryza sativ
6	52	54.7	206	13 Q8QH15	Q8QH15 gallus gall
7	50	52.6	390	5 Q8T8V5	Q8T8V5 drosophila
8	50	52.6	826	10 Q940X9	Q940X9 arabidopsi
9	49	51.6	93	15 Q778B1	Q778B1 human immun
10	49	51.6	170	15 Q91WT4	Q91WT4 human immun
11	48	50.5	90	15 Q69711	Q69711 human immun
12	48	50.5	230	11 Q9D825	Q9D825 mus musculu
13	48	50.5	376	11 Q9ESU0	Q9ESU0 mus musculu
14	48	50.5	385	4 Q9UEH2	Q9UEH2 homo sapien
15	48	50.5	385	11 Q9ESU1	Q9ESU1 mus musculu
16	48	50.5	394	4 Q9H836	Q9H836 homo sapien

17	48	50.5	859	15 O11947	O11947 human immun
18	47	49.5	102	15 Q90AB7	Q90AB7 human immun
19	47	49.5	114	15 Q9YL33	Q9YL33 human immun
20	47	49.5	115	15 Q9YL52	Q9YL52 human immun
21	47	49.5	115	15 Q74722	Q74722 human immun
22	47	49.5	143	15 Q9J6A7	Q9J6A7 human immun
23	47	49.5	343	10 Q94L00	Q94L00 oryza sativ
24	47	49.5	898	11 O88193	O88193 mus musculu
25	47	49.5	934	11 Q91Y15	Q91Y15 mus musculu
26	46	48.4	102	15 Q90AY3	Q90AY3 human immun
27	46	48.4	102	15 Q90AW6	Q90AW6 human immun
28	46	48.4	102	15 Q90AN7	Q90AN7 human immun
29	46	48.4	128	15 Q9QMB7	Q9QMB7 human immun
30	46	48.4	391	2 P72382	P72382 staphylococ
31	46	48.4	391	2 P95709	P95709 staphylococ
32	46	48.4	391	16 Q99X57	Q99X57 staphylococ
33	45	47.4	90	15 Q39383	Q39383 human immun
34	45	47.4	92	15 Q76931	Q76931 human immun
35	45	47.4	92	15 Q76934	Q76934 human immun
36	45	47.4	92	15 Q76937	Q76937 human immun
37	45	47.4	92	15 Q76950	Q76950 human immun
38	45	47.4	92	15 Q76958	Q76958 human immun
39	45	47.4	92	15 Q76959	Q76959 human immun
40	45	47.4	147	15 Q90BB6	Q90BB6 human immun
41	45	47.4	156	16 Q9RZ48	Q9RZ48 deinococcus
42	45	47.4	203	15 Q90004	Q90004 human immun
43	45	47.4	207	15 Q71839	Q71839 human immun
44	45	47.4	213	15 Q90R02	Q90R02 human immun
45	45	47.4	214	15 Q90001	Q90001 human immun

ALIGNMENTS

RESULT 1

O01806 PRELIMINARY; PRT; 396 AA.

ID O01806

AC O01806;

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE C44E4.4 protein.

GN C44E4.4.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RX MEDLINE=99069613; PubMed=9851916;

RA None;

RT "Genome sequence of the nematode C. elegans: a platform for

RT Investigating biology. The C. elegans Sequencing Consortium.";

RL Science 282:2012-2018(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;

RA Sammons L., Wohldmann P., Gillam B.;

RT "The sequence of C. elegans cosmid C44E4.";

RN [3]

RP Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

RC STRAIN-BRISTOL N2;

RA Waterston K.;

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF003140; AAB54169.1; .

DR InterPro; IPR002344; Lupus.La.

DR InterPro; IPR000504; RNA_rec_mot.

DR Pfam; PF00076; rrm; 1.

DR PRINTS; PR00302; LUPUSLA.

DR SMART; SM00360; RRM; 1.

DR PROSITE; PS50102; RRM; 1.

OC eurosids II: Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]

RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL: AY052365; AAK96556.1; -;
DR InterPro: IPR001950; TIF-SUT1.
DR PROSITE: PS01118; SUT1-1; UNKNOWN.1.
SQ SEQUENCE 826 AA; 91377 MW; 8D41922E5B609D9A CRC64;

Query Match 52.6%; Score 50; DB 10; Length 826;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 DADQRIKLEYFGN 17
Db 277 DLRDLKQVEYFSD 292
- - - - -
| : : : : | : : : : |

RESULT 9
Q77881 PRELIMINARY; PRT; 93 AA.
AC Q77881;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Envelope protein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=95191002; PubMed=7884875;
RA Mulder-Kampinga G.A., Simonon A., Kuiken C.L., Dekker J.,
RA Scherpbier H.J., de Perre P., Boer K., Goudsmit J.;
RT "Similarity in env and gag genes between genomic RNAs of human
RT immunodeficiency virus type 1 (HIV-1) from mother and infant is
RT unrelated to time of HIV-1 RNA positivity in the child.";
RL J. Virol. 69:2285-2296(1995).
DR ENBL: Z47894; CAA87908.1; -;
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1 93
FT NON_TER 93 93
SQ SEQUENCE 93 AA; 10561 MW; 4941876B98E5FFA8 CRC64;

Query Match 51.6%; Score 49; DB 15; Length 93;
Best Local Similarity 61.5%; Pred. No. 2;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 5 QRIKLEYFGN 17
Db 73 QKVTQLEYFGN 85
- - - - -
| : : : | : : : |

RESULT 10
Q91WT4 PRELIMINARY; PRT; 170 AA.
ID Q91WT4;
AC Q91WT4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Env glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=21134754; PubMed=11242522;
RA Yang C., Dash B., Hanna S.L., Frances H.S., Nzilambi N.,
RA Colebunders R.C., St Louis M., Quinn T.C., Folks T.M., Lal R.B.;
RT "Predominance of HIV type 1 subtype G among commercial sex workers
RT from Kinshasa, Democratic Republic of Congo.";
RL AIDS Res. Hum. Retroviruses 17:361-365(2001).
DR ENBL: AF260442; AAF71590.1; -;
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1 170
FT NON_TER 170 170
SQ SEQUENCE 170 AA; 19351 MW; 177B41DAF6F7B7E CRC64;

Query Match 51.6%; Score 49; DB 15; Length 170;
Best Local Similarity 52.9%; Pred. No. 3.8;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 DDADQRIKLEYFGN 17
Db 75 DTLQKIVQLRKYFGN 91
- - - - -
| : : : : | : : : : |

RESULT 11
Q69711 PRELIMINARY; PRT; 90 AA.
ID Q69711;
AC Q69711;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Envelope glycoprotein gp120 (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=96392168; PubMed=8798980;
RA Lukashov V.V., Kuiken C.L., Boer K., Goudsmit J.;
RT "Hiv type 1 subtypes in The Netherlands circulating among women
RT originating from AIDS-endemic regions";
RL AIDS Res. Hum. Retroviruses 12:951-953(1996).
DR ENBL: L76905; AAC37965.1; -;
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1 90
FT NON_TER 90 90
SQ SEQUENCE 90 AA; 10247 MW; 42CB50C007F56186 CRC64;

Query Match 50.5%; Score 48; DB 15; Length 90;
Best Local Similarity 69.2%; Pred. No. 2.8;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 5 QRIKLEYFGN 17
Db 73 QKVIKLEYFGN 85
- - - - -
| : : : : | : : : : |

RESULT 12
Q9D825 PRELIMINARY; PRT; 230 AA.
ID Q9D825;
AC Q9D825;
DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE SWI/SNF-related, matrix associated, actin dependent regulator of
 DE chromatin, subfamily b, member 1.
 GN SMARCB1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=SMALL INTESTINE;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Nukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Schoenbach M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schenwald C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK008558; BAB25743.1; -;
 DR MGD; MGI:1328366; Smarcb1.
 SQ SEQUENCE 230 AA; 26444 MW; 71CD44A332D0DB46 CRC64;

 Query Match 50.5%; Score 48; DB 11; Length 230;
 Best Local Similarity 50.0%; Pred. No. 7.6;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

 QY 1 DDADQRIIKQLEYFNGNI 18
 | :|||: : : |||
 DB 101 DQSDQVVIKLNHVNGNI 118

 RESULT 13
 Q9ESU0 ID Q9ESU0 PRELIMINARY; PRT; 376 AA.
 AC Q9ESU0;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DE Integrase interactor.
 DE Integrase interactor 1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H HEN; TISSUE=SPLENOCYTES;
 RA Tozaki H., Yasuda J., Iwakura Y.;
 RT "Cloning of murine Inil 27deletion cDNA."
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB017344; BAB12428.1; -;
 SQ SEQUENCE 376 AA; 43174 MW; 79323C64CCAI4A05 CRC64;

 Query Match 50.5%; Score 48; DB 11; Length 376;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

 QY 1 DDADQRIIKQLEYFNGNI 18
 | :|||: : : |||

Db 247 DQSDQVVIKLNHVNGNI 264.

 RESULT 14
 Q9UBH2 ID Q9UBH2 PRELIMINARY; PRT; 385 AA.
 AC Q9UBH2;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DE Integrase interactor 1A protein (IN11A).
 DE Integrase interactor 1A protein (IN11A).
 GN IN11A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kedra D., Bruder C., Dunanski J.P.;
 RT "Cloning of the mouse integrase interactor 1 gene."
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95258318; PubMed=7739891;
 RA Muchardt C., Sardet C., Bourachot B., Onufryk C., Yaniv M.;
 RT "A human protein with homology to Saccharomyces cerevisiae SNF5
 RT interacts with the potential helicase hbrm."
 RL Nucleic Acids Res. 23:1127-1132(1995).
 DR EMBL; AJ011737; CAA09758.1; -;
 SQ SEQUENCE 385 AA; 44141 MW; B7BCA26875BD943D CRC64;

 Query Match 50.5%; Score 48; DB 4; Length 385;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

 QY 1 DDADQRIIKQLEYFNGNI 18
 | :|||: : : |||
 Db 256 DQSDQVVIKLNHVNGNI 273

 RESULT 15
 Q9ESU1 ID Q9ESU1 PRELIMINARY; PRT; 385 AA.
 AC Q9ESU1;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DE Integrase interactor 1.
 GN IN11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H HEN; TISSUE=SPLENOCYTES;
 RA Tozaki H., Yasuda J., Iwakura Y.;
 RT "Cloning of murine Inil cDNA."
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB017343; BAB12427.1; -;
 SQ SEQUENCE 385 AA; 44157 MW; A2FCADE642DFB991 CRC64;

 Query Match 50.5%; Score 48; DB 11; Length 385;
 Best Local Similarity 50.0%; Pred. No. 13;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

 QY 1 DDADQRIIKQLEYFNGNI 18
 | :|||: : : |||
 Db 256 DQSDQVVIKLNHVNGNI 273

 Search completed: April 23, 2003, 13:32:57
 Job time : 29.9101 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:42 ; Search time 4.95506 Seconds
(without alignments)
150.669 Million cell updates/sec

Title: US-09-836-073-17

Perfect score: 95
Sequence: 1 DDADQRIIKLEYFGNI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	60.0	428	LAA_XENLA	P28048 xenopus lae
2	55	57.9	427	LAB_XENLA	P28049 xenopus lae
3	50	52.6	390	LA_DROME	P40796 drosophila
4	48	50.5	385	1 SNF5_HUMAN	Q12824 homo sapien
5	48	50.5	385	1 SNF5_MOUSE	Q920h3 mus musculu
6	48	50.5	936	1 CDA5_HUMAN	Q9y5h7 homo sapien
7	47	49.5	383	1 LA_AEDAL	Q26457 aedes albop
8	45	47.4	177	1 IPYR_NEIMA	Q91vg3 neisseria m
9	45	47.4	177	1 IPYR_NEIMB	Q9kg94 neisseria m
10	45	47.4	1002	1 YEMA_DROME	P25992 drosophila
11	44	46.3	404	1 LA_BOVIN	P10881 bos taurus
12	44	46.3	408	1 LA_HUMAN	P05455 homo sapien
13	44	46.3	415	1 LA_MOUSE	P32067 mus musculu
14	44	46.3	415	1 LA_RAT	P38556 rattus norv
15	44	46.3	2283	1 DPOE_MOUSE	Q9wxf7 mus musculu
16	43	45.3	176	1 IPYR_HAEIN	P44529 haemophilus
17	43	45.3	1290	1 RPOC_MYCPN	P75271 mycoplasma
18	42	44.2	214	1 EQ4_ACTEQ	Q9y1u9 actinia equ
19	42	44.2	222	1 KADC_MAIZE	P43188 zea mays (m
20	42	44.2	365	1 ROAL_DROME	P07909 drosophila
21	42	44.2	880	1 CADF_XENLA	P33148 xenopus lae
22	42	44.2	884	1 CADB_XENLA	P33152 xenopus lae
23	41.5	43.7	729	1 KEX1_YEAST	P09620 saccharomyc
24	41	43.2	178	1 COAD_ZYMMO	Q9rme4 zymomonas m
25	41	43.2	211	1 MT04_SYNY3	P73161 synchocyst
26	41	43.2	214	1 KAD_MYCPU	Q98q02 mycoplasma
27	41	43.2	298	1 LAHL_SCHPO	P87058 schizosacch
28	41	43.2	315	1 HO2_RAT	P23711 rattus norv
29	41	43.2	361	1 YPDF_ECOLI	P76524 escherichia
30	41	43.2	1170	1 SMC2_YEAST	P38989 saccharomyc
31	41	43.2	1292	1 RPOC_MYCGE	P47582 mycoplasma
32	40	42.1	108	1 Y165_UREPA	Q9p9x9 ureaplasma
33	40	42.1	594	1 SYA_BORBU	O51238 borrelia bu

RESULT 1

ID	LAA_XENLA	STANDARD;	PRT;	428 AA.
AC	P28048;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Lupus la protein homolog A (La ribonucleoprotein A) (La autoantigen homolog A).			
GN	LAA1.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-Oocyte;			
RX	MEDLINE=93287095; PubMed=8510143;			
RA	Scherly D., Stutz F., Lin-Marq N., Clarkson S.G.;			
RT	"La proteins from Xenopus laevis. cDNA cloning and developmental			
RT	expression."			
RL	J. Mol. Biol. 231:196-204(1993).			
CC	-1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA			
CC	POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION			
CC	FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT			
CC	POLYMERASE III TRANSCRIPTS (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: Nuclear (Probable).			
CC	-1- DEVELOPMENTAL STAGE: BARELY DETECTABLE IN STAGE I/II OOCYTES,			
CC	ACCUMULATE IN STAGE III/IV OOCYTES, THEN EXHIBIT A ROUGHLY			
CC	CONSTANT STEADY STATE LEVEL IN MATURE OOCYTES, EGGS, AND EARLY			
CC	EMBRYOS.			
CC	-1- PTM: PHOSPHORYLATED (PROBABLE).			
CC	-1- MISCELLANEOUS: THERE ARE TWO FORMS OF LA, LAA AND LAB, IN XENOPUS.			
CC	-1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.			
CC	-1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/).			
CC	or send an email to license@isb-sib.ch .			
CC	-----			
DR	EMBL; X68817; CAA48715.1; -			
DR	PIR; S28545; S28545.			
DR	PIR; S33818; S33818.			
DR	InterPro: IPR002344; Lupus_La.			
DR	InterPro: IPR000504; RNA_rec_mot.			
DR	Pfam; PF000076; rrm; 1.			
DR	PRINTS; PR00302; LUPUSLA.			
DR	SMART; SM00360; RRM; 1.			
DR	PROSITE; PS0102; RRM; 1.			
DR	PROSITE; PS00030; RRM_RNP_1; 1.			
KW	RNA-binding; Nuclear protein; Phosphorylation.			

O54215 streptomyc
P43075 candida alb
P53114 saccharomyc
P18062 cyanophora
P14960 quillardia
P26575 solanum tub
P26576 solanum tub
P26577 solanum tub
P10647 solanum tub
P32764 solanum tub
P57913 pasteurella
O83296 treponema p

ALIGNMENTS

```

FT DOMAIN      111          203         RNA-BINDING (RRM).
FT DOMAIN      196          212        NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN      316          332        NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE    428 AA; 48864 MW; AEB3A3B87D2E3EC3 CRC64;

Query Match           60.0%; Score 57; DB 1; Length 428;
Best Local Similarity 56.2%; Pred. No. 0.055;
Matches     9; Conservative     5; Mismatches     2; Indels   0; Gaps   0;

QY       2 DADQRRIKOLEYYFGN 17
DB       12 DSTKTCEQIEYFGD 27
| : | : | : | : | : | : | : | :
|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|-|

RESULT 2
LAB_XENLA              STANDARD;             PRT;      427 AA.
ID LAB_XENLA
AC P28049;
DT DT
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lups La protein homolog B (La ribonucleoprotein B) (La autoantigen
DE homolog B).
DE DE
GN LABL.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oocyte;
RX MEDLINE=93387095; PubMed=8510143;
RA Scherly D. Stutz F., Lin-Marq N., Clarkson S.G.;
RT "La proteins from xenopus laevis. cDNA cloning and developmental
expression."; RT
RL J. Mol. Biol. 231:196-204(1993).
CC -I- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Nuclear (probable).
CC -I- DEVELOPMENTAL STAGE: BARELY DETECTABLE IN STAGE I/II OOCYTES,
CC ACCUMULATE IN STAGE III/IV OOCYTES, THEN EXHIBIT A ROUGHLY
CC CONSTANT STEADY STATE LEVEL IN MATURE OOCYTES, EGGS, AND EARLY
CC EMBRYOS
CC -I- PTM: PHOSPHORYLATED (PROBABLE).
CC -I- MISCELLANEOUS: THERE ARE TWO FORMS OF LA, LAA AND LAB, IN XENOPUS.
CC -I- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
CC -I- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
-----
This SWISS-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation at
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X68818; CAA48716.1; -.
DR DR
DR PIR; S28544; S28544.
DR DR
DR PIR; S33817; S33817.
DR DR
DR InterPro; IPRO02344; Lupus_La.
DR DR
DR InterPro; IPRO00504; RNa_rec_mot.
DR DR
DR Pfam; PF00076; rrm; 1.
DR DR
DR PRINTS; PR00302; LUPUSLA.
DR DR
DR SMART; SM00360; RRM; 1.
DR DR
DR PROSITE; PS0102; RRM; 1.
DR DR
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
RNA-binding; Nuclear protein; Phosphorylation.
FT DOMAIN      110          202
FT DOMAIN      315          331        RNA-BINDING (RRM).
SQ SEQUENCE    427 AA; 448995 MW; 45F3146F8934A355 CRC64;

```

```

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT The genome sequence of Drosophila melanogaster.
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTION TERMINATION BY RNA
CC POLYMERASE III. BINDS RNA AND DNA. BINDS TO PRECURSORS OF RNA
CC POLYMERASE III TRANSCRIPTS. MAY PLAY A SPECIALIZED ROLE DURING FLY
CC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT EMBRYONIC, LARVAL,
CC PUPAL, AND ADULT DEVELOPMENT. EXPRESSION THROUGHOUT THE EMBRYO IS
CC FOLLOWED BY A RESTRICTED PATTERN OF MESODERMAL EXPRESSION THAT IS
CC LATER CONFINED TO THE VISCERAL MESODERM, GONADS, GUT, AND SALIVARY
CC GLANDS.
CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -!- SIMILARITY: TO VERTEBRATE PROTEIN LA.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U07652; AAA20518.1; -.
CC DR EMBL; L32988; AAA21776.1; -.
CC DR EMBL; AE003666; AAF53885.1; -.
CC DR FlyBase; FBgn0011638; La.
CC DR InterPro; IPR002344; Lupus_La.
CC DR Pfam; PF00076; rrm; 1.
CC DR SMART; PR00302; LUPUSLA.
CC DR PRINTS; PS0102; RRM; 1.
CC DR PROSITE; PS00030; RRM_RNP_1; 1.
CC DR RNA-binding; Nuclear protein; DNA-binding.
CC KW DOMAIN 149 234 RNA-BINDING (RRM).
CC FT CONFLICT 169 169 A -> T (IN REF. 1).
CC FT CONFLICT 182 183 KH -> NS (IN REF. 1).
CC FT CONFLICT 283 283 A -> R (IN REF. 1).
CC FT CONFLICT 329 329 K -> N (IN REF. 1).
CC SQ SEQUENCE 390 AA; 44884 MW; A8099288B90446A5 CRC64;

Query Match 52.6%; Score 50; DB 1; Length 390;
Best Local Similarity 57.1%; Pred. No. 0.74;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 4 DORIINKLEYEYFGN 17
Db 51 ERAIRQVEYFEGD 64

RESULT 4
SNF5_HUMAN STANDARD; PRT; 385 AA.
AC Q12824; O95474; O75784;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE SWI/SNF related, matrix associated, actin dependent regulator of
DE chromatin subfamily B member 1 (Integrase Interactor 1 protein)
DE (hSNF5) (BAF47).
GN SMARCB1 OR SNF5L1 OR INI1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95099327; PubMed=7801128;
RA Kalpana G.V., Marmon S., Wang W., Crabtree G.R., Goff S.P.;
RT Binding and stimulation of HIV-1 integrase by a human homolog of
RT yeast transcription factor SNF5.
RL Science 286:2002-2006(1994).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=9925325; PubMed=10208979;
RA Bruder C.E., Dumanski J.P., Kedra D.;
RT "The mouse ortholog of the human SMARCB1 gene encodes two splice
RT forms."
RL Biochem. Biophys. Res. Commun. 257:886-890(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98334382; PubMed=9671307;
RA Versteeg I., Sevenet N., Lange J., Rousseau-Merck M.F., Ambros P.,
RA Handgretinger R., Aurias A., Delattre O.;
RT "Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer."
RL Nature 394:203-206(1998).
CC -!- FUNCTION: INVOLVED IN CHROMATIN-REMODELING. PART OF A COMPLEX THAT
CC OPENS THE CHROMATIN TO FACILITATE THE TRANSCRIPTIONAL MACHINERY TO
CC ACCESS THEIR TARGETS.
CC -!- SUBUNIT: BINDS TIGHTLY TO THE HUMAN IMMUNODEFICIENCY VIRUS-TYPE 1
CC (HIV-1) INTEGRASE IN VITRO AND STIMULATES ITS DNA-JOINING
CC ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A/INI1A (SHOWN HERE) AND
CC B/INI1B; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- DISEASE: TUMOR SUPPRESSOR. INACTIVATED IN MALIGNANT RHABDOID
CC TUMORS (MRT). MRT'S ARE EXTREMELY AGGRESSIVE CANCERS OF EARLY
CC CHILDHOOD.
CC -!- SIMILARITY: BELONGS TO THE SNF5 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U04847; AAA81905.1; -.
CC DR EMBL; AJ011738; CAA09759.1; -.
CC DR EMBL; Y17118; CAA76639.1; ALT_SEQ.
CC DR EMBL; Y17119; CAA76639.1; JOINED.
CC DR EMBL; Y17120; CAA76639.1; JOINED.
CC DR EMBL; Y17121; CAA76639.1; JOINED.
CC DR EMBL; Y17122; CAA76639.1; JOINED.
CC DR EMBL; Y17123; CAA76639.1; JOINED.
CC DR EMBL; Y17124; CAA76639.1; JOINED.
CC DR EMBL; Y17125; CAA76639.1; JOINED.
CC DR EMBL; Y17126; CAA76639.1; JOINED.
CC DR MIM; 601607; -.
CC Transcription regulation; Activator; Nuclear protein;
KW Alternative splicing; Anti-oncogene.
FT VARSPLIC 69 77 MISSING (IN ISOFORM B).
FT CONFLICT 136 136 S -> P (IN REF. 2).
FT CONFLICT 378 378 L -> E (IN REF. 3).
FT CONFLICT 382 382 G -> A (IN REF. 2).
FT CONFLICT 385 385 B7BC6D6DD5BC443D CRC64;
SQ SEQUENCE 385 AA; 44117 MW; B7BC6D6DD5BC443D CRC64;

Query Match 50.5%; Score 48; DB 1; Length 385;
Best Local Similarity 50.0%; Pred. No. 1.6;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIKQLEYEYFGN 18
Db 256 DQSDQRVVIKLIHVGNI 273

```

```
RESULT 5
SNF5_MOUSE STANDARD; PRT; 385 AA.
ID AC Q920H3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE SWI/SNF related, matrix associated, actin dependent regulator of
DE chromatin subfamily B member 1 (Integrase interactor 1 protein)
DE (SNF5)
DE SNARCB1 OR SNF5L1 OR INI1.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=99225325; PubMed=10208879;
RA Bruder C.E., Dumanski J.P., Kedra D.;
RT "The mouse ortholog of the human SNARCB1 gene encodes two splice
RT forms";
RL Biochem. Biophys. Res. Commun. 257:886-890(1999).
CC -1- FUNCTION: INVOLVED IN CHROMATIN-REMODELING. PART OF A COMPLEX THAT
CC OPENS THE CHROMATIN TO FACILITATE THE TRANSCRIPTIONAL MACHINERY TO
CC ACCESS THEIR TARGETS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A/INI1A (SHOWN HERE) AND
CC B/INI1B; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE SNF5 FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ011740; CAA09761.1; -
CC EMBL; AJ011739; CAA09760.1; -
CC MGD; MGI:1328366; Smarcb1.
KW Transcription regulation; Activator; Nuclear protein;
KW Alternative splicing.
FT VARSPLIC 69 77 MISSING (IN ISOFORM B).
SQ SEQUENCE 385 AA; 44141 MW; 87BCA26875BD943D CRC64;
Query Match 50.5%; Score 48; DB 1; Length 385;
Best Local Similarity 50.0%; Pred. No. 1.6;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 DDADQRIKOLEYYFGNI 18
Db 256 DQSDQRVILNIHVGN 273
RESULT 6
CDA5_HUMAN STANDARD; PRT; 936 AA.
ID AC Q9Y5H7; O75284;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protocadherin alpha 5 precursor (PCDH-alpha5).
GN PCDH5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RX MEDLINE=99308636; PubMed=10380929;
```

```
RA Wu Q., Maniatis T.;
RT "A striking organization of a large family of human neural cadherin-
RL like cell adhesion genes.";
RN Cell 97:779-790(1999).
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Kimmerly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,
RA Kader K., Miguel T., Miller C., Pittluck S., Poillard M., Rojeski H.,
RA Subramanian S., Martin C.H.;
RT "Sequencing of human chromosome 5";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: POTENTIAL CALCIUM-DEPENDENT CELL-ADHESION PROTEIN.
CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein (by similarity).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF152313; AAD43707.1; -
CC EMBL; AF152483; AAD43744.1; -
CC EMBL; AC005609; AAC34321.1; -
CC Genew; HGNC:8671; PCDH5.
DR MIM; 606311; -
DR MIM; 604966; -
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 5.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 6.
DR PROSITE; PS00232; CADHERIN_1; 5.
DR PROSITE; PS00268; CADHERIN_2; 6.
KW Calcium-binding; Cell adhesion; Glycoprotein; Signal; Repeat;
KW Transmembrane; Alternative splicing; Multigene family.
FT SIGNAL 1 28
FT CHAIN 29 936
FT DOMAIN 29 696
FT TRANSMEM 697 717
FT DOMAIN 718 936
FT DOMAIN 33 132
FT DOMAIN 156 241
FT DOMAIN 242 349
FT DOMAIN 350 454
FT DOMAIN 455 564
FT DOMAIN 580 677
FT DOMAIN 909 916
FT DOMAIN 773 890
FT REPEAT 773 776
FT REPEAT 785 788
FT REPEAT 818 821
FT REPEAT 873 876
FT REPEAT 877 890
FT CARBOHYD 264 264
FT CARBOHYD 547 547
FT VARSPLIC 785 816
FT VARSPLIC 817 936
FT SEQUENCE 936 AA; 102047 MW; 7143EECE1A357CD6 CRC64;
Query Match 50.5%; Score 48; DB 1; Length 936;
Best Local Similarity 47.1%; Pred. No. 4.3;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 2 DADQRIKOLEYYFGNI 18
Db 267 DADEGINKEIVYFNSL 283
```

```

RESULT 7
LA_AEDAL STANDARD; PRT; 383 AA.
ID LA_AEDAL
AC Q26457;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
OS Aedes albopictus (Forest day mosquito).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Culicoidae; Aedes.
OX NCBI_TaxID=7160;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96135233; PubMed=8551578;
RA Pardigon N., Strauss J.H.;
RT "Mosquito homolog of the La autoantigen binds to Sindbis virus RNA.";
RL J. Virol. 70:1173-1181(1996).
CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTION TERMINATION BY RNA
CC POLYMERASE III. BINDS RNA AND DNA. BINDS TO THE 3' END OF THE
CC MINUS STRAND OF SINDBIS VIRUS RNA. THIS MAY BE SIGNIFICANT FOR
CC SINDBIS VIRUS RNA REPLICATION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. PRIMARILY NUCLEAR,, BUT SIGNIFICANT
CC AMOUNTS ARE PRESENT IN THE CYTOPLASM.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S80954; AAB35931.1; -
CC InterPro; IPR002344; Lupus.La.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS00102; RRM; 1.
CC PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
CC RNA-binding; Nuclear protein; DNA-binding.
KW DOMAIN 141 228
FT DOMAIN 141 228 RNA-BINDING (RRM).
SQ SEQUENCE 383 AA; 44430 MW; 4E5CC8F21C40F452 CRC64;

Query Match 49.5%; Score 47; DB 1; Length 383;
Best Local Similarity 80.0%; Pred. No. 2.3;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 8 IKQLEYFGN 17
I:|||||:
Db 48 IQLEYFGD 57

RESULT 8
IPYR_NEIMA STANDARD; PRT; 177 AA.
ID IPYR_NEIMA
AC Q9JVG3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
DE hydrolyase) (Ppase).
GN PPA OR NMA0851.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leathes S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491.";
RL Nature 404:502-506(2000).
CC -1- CATALYTIC ACTIVITY: Diphosphate + H(2)O -> 2 phosphate.
CC -1- COFACTOR: Requires the presence of divalent metal cation.
CC Magnesium confers the highest activity. Binds 4 divalent cations
CC per subunit (By similarity).
CC -1- SUBUNIT: Homohexamer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE PPASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL162754; CAB84132.1; -
CC HSSP; P17288; 1FAJ.
CC InterPro; IPR001596; Pyrophosphatase.
CC Pfam; PF00719; Pyrophosphatase; 1.
CC PRODOM; PD002014; Pyrophosphatase; 1.
CC PROSITE; PS00387; PPASE; 1.
CC Hydrolyase; Magnesium; Complete proteome.
FT ACT_SITE 31 31 BY SIMILARITY.
SQ SEQUENCE 177 AA; 19799 MW; 1DAB8BF80F8FC06F CRC64;

Query Match 47.4%; Score 45; DB 1; Length 177;
Best Local Similarity 43.8%; Pred. No. 2.1;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Oy 2 DADQRIIKQLEYFGN 17
I:|||||:
Db 126 DLPQQLIKQIEFHFNH 141

RESULT 9
IPYR_NEIMB STANDARD; PRT; 177 AA.
ID IPYR_NEIMB
AC Q9K0G4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
DE hydrolyase) (Ppase).
GN PPA OR NMB0641.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,
RA Hart D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";

```

RL Science 287:1809-1815(2000).
 CC 1- CATALYTIC ACTIVITY: Diphosphate + H(2)O -> 2 phosphate.
 CC 1- COFACTOR: Requires the presence of divalent metal cation.
 CC MAGNESIUM confers the highest activity. Binds 4 divalent cations
 CC per subunit (By similarity).
 CC 1- SUBUNIT: Homohexamer (By similarity).
 CC 1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC 1- SIMILARITY: BELONGS TO THE PPAASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF002419; AAF41064.1;
 CC HSSP; P17288; 1FAJ.
 CC TIGR; NMB0641;
 CC InterPro: IPR001596; Pyrophosphatase.
 CC Pfam: PF00719; Pyrophosphatase; 1.
 CC ProDom: PD002014; Pyrophosphatase; 1.
 CC PROSITE; PS00387; PPAASE; 1.
 CC Hydrolase; Magnesium; Complete proteome.
 CC ACT_SITE 31 31 BY SIMILARITY.
 CC SEQUENCE 177 AA; 19811 MW; 6017182C446A8567 CRC64;
 CC
 CC Query Match 47.4%; Score 45; DB 1; Length 177;
 CC Best Local Similarity 43.8%; Pred. No. 2.1;
 CC Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 2 DADQRIKOLEYVFGN 17
 CC |:::||||:|:
 CC Db 126 DLPOQLIKQIEFHNH 141
 CC
 CC RESULT 10
 CC YEMA DROME
 CC ID YEMA DROME STANDARD; PRT; 1002 AA.
 CC AC P25992; Q9VAP4;
 CC DT 01-MAY-1992 (Rel. 22, Created)
 CC DT 01-MAY-1992 (Rel. 22, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Yemanuclein-alpha.
 CC GN YEMA-ALPHA OR YEMA OR YG4.5 OR CG11879.
 CC OS Drosophila melanogaster (Fruit fly).
 CC OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 CC OX NCBI_TaxID=7227;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=Canton-S;
 CC RX MEDLINE=92297435; PubMed=1606021;
 CC RA Alt-Ahmed O., Bellon B., Capri M., Joblet C., Thomas-Delaage M.;
 CC RT "The yemanuclein-alpha: a new Drosophila DNA binding protein specific
 CC for the oocyte nucleus.";
 CC RL Mech. Dev. 37:69-80(1992).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=Berkely;
 CC RX MEDLINE=20196006; PubMed=10731132;
 CC RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 CC RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 CC RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 CC RA Sutton G.G., Wortman J.R., Vandeell M.D., Zhang O., Chen L.X.,
 CC RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 CC RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.D.G.,
 CC RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 CC RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 CC RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 CC RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dougan K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhang W., Zhou X., Zhou Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC 1- FUNCTION: MAY PLAY A KEY ROLE IN EGG ORGANIZATION. IT MAY BE A
 CC TRANSCRIPTIONAL REGULATOR.
 CC 1- SUBCELLULAR LOCATION: Nuclear.
 CC 1- TISSUE SPECIFICITY: OOCYTE-SPECIFIC.
 CC 1- DEVELOPMENTAL STAGE: EXPRESSED AT ALL OOCYtic STAGES.
 CC 1- PTM: THE N-TERMINUS IS BLOCKED.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X63503; CAA45074.1;
 CC EMBL; AE003768; AAF56858.2;
 CC PIR; S22146; S22146.
 CC Flybase; FBgn0005596; yem-alpha.
 CC KW Nuclear protein; DNA-binding; Repeat.
 CC FT DOMAIN 80 85 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC FT DOMAIN 207 217 POLY-SER.
 CC FT DOMAIN 219 261 ASP/GLU-RICH (ACIDIC).
 CC FT DOMAIN 230 253 2 X 12 AA TANDEM REPEATS.
 CC FT REPEAT 230 241 1.
 CC FT REPEAT 242 253 2.
 CC FT VARIANT 698 698 S -> L.
 CC SQ SEQUENCE 1002 AA; 109310 MW; EE69A384EBA24D2F CRC64;
 CC
 CC Query Match 47.4%; Score 45; DB 1; Length 1002;
 CC Best Local Similarity 52.9%; Pred. No. 15;
 CC Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 CC
 CC QY 1 DDADQRIKOLEYVFGN 17
 CC |:::||||:|:
 CC Db 105 DDDVARIVRELEAKYGN 121
 CC
 CC RESULT 11
 CC LA_BOVIN
 CC ID LA_BOVIN STANDARD; PRT; 404 AA.
 CC AC P10881;
 CC DT 01-JUL-1989 (Rel. 11, Created)
 CC DT 01-JUL-1989 (Rel. 11, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)


```

QY 4 DQRIKOLEYFNG 17
Db 14 EAKICHOIEYFGD 27

RESULT 15
DPOE_MOUSE STANDARD; PRT; 2283 AA.
AC Q9WVF7; Q9QX50;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase epsilon, catalytic subunit A (EC 2.7.7.7) (DNA
  polymerase II subunit A).
GN POLE OR POLE1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99296371; PubMed=10366722;
RA Huang D., Knuuti R., Palosaari H., Pospiech H., Syvaioja J.E.;
RT "cDNA and structural organization of the gene Pole1 for the mouse DNA
  polymerase epsilon catalytic subunit.";
RL Biochim. Biophys. Acta 1445:363-371(1999).
CC -!- FUNCTION: PARTICIPATES IN DNA REPAIR AND IN CHROMOSOMAL DNA
  REPLICATION.
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
  + [DNA](N).
CC -!- SUBUNIT: CONSISTS OF TWO SUBUNITS (258 kDa AND 55 kDa).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: THE DNA POLYMERASE ACTIVITY DOMAIN RESIDES IN THE
  N-TERMINAL HALF OF THE PROTEIN, WHILE THE C-TERMINUS IS NECESSARY
  FOR COMPLEXING SUBUNITS B AND C. THE C-TERMINUS MAY ALSO REGULATE
  THE CATALYTIC ACTIVITIES OF THE ENZYME.
CC -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY. HIGH
  SIMILARITY WITH FUNGAL DNA POLYMERASE II.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AF123502; AAD45244.1;
DR EMBL; AF126398; AAD46482.1;
DR EMBL; AF126377; AAD46482.1; JOINED.
DR EMBL; AF126378; AAD46482.1; JOINED.
DR EMBL; AF126379; AAD46482.1; JOINED.
DR EMBL; AF126380; AAD46482.1; JOINED.
DR EMBL; AF126381; AAD46482.1; JOINED.
DR EMBL; AF126382; AAD46482.1; JOINED.
DR EMBL; AF126383; AAD46482.1; JOINED.
DR EMBL; AF126384; AAD46482.1; JOINED.
DR EMBL; AF126385; AAD46482.1; JOINED.
DR EMBL; AF126386; AAD46482.1; JOINED.
DR EMBL; AF126387; AAD46482.1; JOINED.
DR EMBL; AF126388; AAD46482.1; JOINED.
DR EMBL; AF126389; AAD46482.1; JOINED.
DR EMBL; AF126390; AAD46482.1; JOINED.
DR EMBL; AF126391; AAD46482.1; JOINED.
DR EMBL; AF126392; AAD46482.1; JOINED.
DR EMBL; AF126393; AAD46482.1; JOINED.
DR EMBL; AF126394; AAD46482.1; JOINED.
DR EMBL; AF126395; AAD46482.1; JOINED.
DR EMBL; AF126396; AAD46482.1; JOINED.
DR EMBL; AF126397; AAD46482.1; JOINED.
DR MGD; MGI:1196391; Pole.
InterPro: IPR002064; DNA_pol_B.

```

```

DR Pfam; PF00136; DNA_pol_B; 1.
DR SMART; SM00486; POLBc; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication; DNA repair;
KW DNA-binding; Zinc-finger; Nuclear protein.
FT ZN_FING 2155 2235 POTENTIAL.
FT CONFLICT 205 205 D -> E (IN REF. 1; AAD45244).
FT CONFLICT 284 284 T -> K (IN REF. 1; AAD45244).
FT CONFLICT 1309 1309 G -> E (IN REF. 1; AAD45244).
FT CONFLICT 1320 1320 K -> R (IN REF. 1; AAD45244).
FT CONFLICT 1386 1388 ALP -> GLEF (IN REF. 1; AAD45244).
FT CONFLICT 1665 1665 T -> I (IN REF. 1; AAD45244).
FT CONFLICT 2086 2086 I -> V (IN REF. 1; AAD45244).
FT CONFLICT 2111 2111 L -> O (IN REF. 1; AAD45244).
SQ SEQUENCE 2283 AA; 262025 MW; 7005C34354E465F3 CRC64;

Query Match 46.3%; Score 44; DB 1; Length 2283;
Best Local Similarity 42.9%; Pred. No. 55;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 DADQRIKOLEYF 15
Db 73 DEKRLVSVDYF 86

Search completed: April 23, 2003, 13:28:11
Job time : 5.95506 secs

```

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:03 ; Search time 10.1124 Seconds
(without alignments)
171.119 Million cell updates/sec

Title: US-09-836-073-17
Perfect score: 95
Sequence: 1 DDADQRIIKQLEYFGNI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	92	96.8	396	2	T30953	hypothetical prote
2	57	60.0	428	1	S33818	ribonucleoprotein
3	55	57.9	427	1	S33817	ribonucleoprotein
4	54	56.8	529	2	T00677	hypothetical prote
5	50	52.6	390	2	A33773	La/SS-B homolog D-
6	50	52.6	390	2	A33781	ribonucleoprotein
7	48	50.5	385	2	S54705	transcription fact
8	46	48.4	391	2	H97777	capsular polysacch
9	45	47.4	156	2	D75605	hypothetical prote
10	45	47.4	177	2	F81175	inorganic pyrophos
11	45	47.4	177	2	G81930	probable inorganic
12	45	47.4	1002	2	A56678	yeamanucin-alpha
13	44.5	46.8	823	2	F64526	adenine/cytosine D
14	44	46.3	158	2	AD2410	hypothetical prote
15	44	46.3	300	2	A13101	transcription regu
16	44	46.3	332	2	A96185	probable transcrip
17	44	46.3	334	2	G97972	conserved hypoteth
18	44	46.3	404	1	S03849	ribonucleoprotein
19	44	46.3	408	1	A31888	ribonucleoprotein
20	44	46.3	415	1	JC1494	ribonucleoprotein
21	44	46.3	455	2	G95104	hypothetical prote
22	43	45.3	176	2	G84049	inorganic pyrophos
23	43	45.3	342	2	A70183	hypothetical prote
24	43	45.3	780	2	T27669	hypothetical prote
25	43	45.3	1290	2	S73653	DNA-directed RNA p
26	42	44.2	183	2	F90570	hypothetical prote
27	42	44.2	222	2	S45634	adenylate kinase (
28	42	44.2	340	2	A99798	hypothetical prote
29	42	44.2	365	2	A26459	helix-destabilizin

30	42	44.2	840	2	B87467	conserved hypothet
31	42	44.2	895	1	IXLCP	EP-cadherin precur
32	42	44.2	905	2	S43064	cadherin - African
33	41.5	43.7	188	2	E97104	phospholipase D fa
34	41.5	43.7	729	2	A29651	KEX1 protein precu
35	41	43.2	178	2	S77850	probable histidine
36	41	43.2	192	1	S75273	hypothetical prote
37	41	43.2	214	2	G90582	adenylate kinase (
38	41	43.2	298	2	T38937	rna binding protei
39	41	43.2	298	2	T43542	RNA-binding protei
40	41	43.2	315	1	A35199	heme oxygenase (de
41	41	43.2	346	2	S76647	hypothetical prote
42	41	43.2	350	2	G90199	acyl carrier prote
43	41	43.2	361	1	F85012	hypothetical prote
44	41	43.2	434	2	AG0512	puatative sulfatas
45	41	43.2	450	2	T21931	hypothetical prote

ALIGNMENTS

RESULT 1

T30953
hypothetical protein C44E4.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
C:Accession: T30953
R:Sammons, L.; Wohldmann, P.; Gillam, B.
submitted to the EMBL Data Library, August 1999
A:Description: The sequence of C. elegans cosmid C44E4.
A:Reference number: 220945
A:Accession: T30953
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-396 <SAMS>
A:Cross-references: EMBL:AF003140; PIDN:AAB54169.1
A:Experimental source: strain Bristol N2; clone C44E4
C:Genetics:
A:Map position: 1
A:Introns: 45/1; 114/3
A>Note: C44E4.4
C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat_homology

Query Match 96.8%; Score 92; DB 2; Length 396;
Best Local Similarity 94.4%; Pred. No. 2.7e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DDADQRIIKQLEYFGNI 18
|||||:|||||:|||||
DB 10 DDADQRIIKQLEYFGNI 27

RESULT 2

S33818
ribonucleoprotein La.A - African clawed frog
N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C:Species: Xenopus laevis (African clawed frog)
C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: S33818; S28545
R:Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.
J. Mol. Biol. 231, 196-204, 1993
A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression
A:Reference number: S33817; MUID:93287095; PMID:8510143
A:Accession: S33818
A:Molecule type: mRNA
A:Residues: 1-428 <SCH>
A:Cross-references: EMBL:X68817; NID:964873; PIDN:CAA48715.1; PID:964874
C:Comment: This protein associates with a variety of small RNA molecules, most of
ay act as a transcription termination factor.
C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat_homology
C:Keywords: phosphoprotein; RNA binding
F:112-178/Domain: Ribonucleoprotein repeat_homology <RRM>
F:113-118/Region: RNA-binding RNP2 motif


```

RESULT 7
S54705
transcription factor SNF5 homolog INI1 - human
N:Alternate names: Integrase interactor 1
C:Species: Homo sapiens (man)
C:Date: 27-Oct-1995 #sequence_revision 10-Nov-1995 #text_change 17-Mar-1999
C:Accession: S54705; A55388
R:Muchardt, C.; Sardet, C.; Bourachot, B.; Onufryk, C.; Yaniv, M.
Nucleic Acids Res. 23, 1127-1132, 1995
A:Title: A human protein with homology to Saccharomyces cerevisiae SNF5 interacts with
A:Reference number: S54705; MUID:95258318; PMID:7739891
A:Accession: S54705
A:Molecule type: mRNA
A:Residues: 1-385 <MUC>
A>Note: the sequence from Fig. 1 is inconsistent with that from Fig. 2 in lacking 228-Le
R:Kalpala, G.V.; Marmon, S.; Wang, W.; Crabtree, G.R.; Goff, S.P.
Science 266, 2002-2006, 1994
A:Title: Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcri
A:Reference number: A55388; MUID:95099327; PMID:7801128
A:Accession: A55388
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-135, 'S', 137-385 <KAL>
A:Cross-references: GB:U04847
A:Experimental source: cell line HL60
C:Genetics:
A:Gene: GDB:SNF5L1; INI1
A:Map position: 3q25.1-3q26.1
C:Function:
A:Description: DNA binding; transcription activation; stimulates human HIV integrase act
C:Keywords: transcription regulation

Query Match 50.5%; Score 48; DB 2; Length 385;
Best Local Similarity 50.0%; Pred. No. 4.7;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIIKOLEYFGNI 18
|:||||:|:|:|
Db 256 DQSDQRIIKLHVGN 273

RESULT 8
H89777
capsular polysaccharide synthesis enzyme Cap5P [imported] - Staphylococcus aureus (strain
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: H89777
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: H89777
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <KUR>
A:Cross-references: GB:BA000018; PID:g13700080; PIDN:BA841379.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: capP
C:Superfamily: lipopolysaccharide biosynthesis protein bplD

Query Match 48.4%; Score 46; DB 2; Length 391;
Best Local Similarity 57.1%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 AQDRIIKOLEYFG 16
|:|:|:|:|
Db 363 ASRRICEAIEYFG 376

```

```

RESULT 9
D75605
hypothetical protein - Deinococcus radiodurans (strain RL)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: D75605
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson,
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski,
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: D75605
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-156 <WHI>
A:Cross-references: GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AAF12329.1; PID:g
A:Experimental source: strain RL
C:Genetics:
A:Gene: DRA0106
A:Map position: 2

Query Match 47.4%; Score 45; DB 2; Length 156;
Best Local Similarity 47.1%; Pred. No. 5.6;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIIKOLEYFGN 17
|:||||:|:|:|
Db 91 EDADQATALNQLQYAVN 107

RESULT 10
F81175
inorganic pyrophosphatase NMB0641 [imported] - Neisseria meningitidis (strain MC58
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: F81175
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Else
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty,
Ric, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli,
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC5
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: F81175
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-177 <TET>
A:Cross-references: GB:AE002419; GB:AE002098; NID:g7225863; PIDN:AAF41064.1; PID:g
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB0641
C:Superfamily: inorganic pyrophosphatase

Query Match 47.4%; Score 45; DB 2; Length 177;
Best Local Similarity 43.8%; Pred. No. 6.4;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 DADQRIIKOLEYFGN 17
|:||||:|:|:|
Db 126 DLPQQLIKQIEFHFNH 141

RESULT 11
G81930
probable inorganic diphosphatase (EC 3.6.1.1) NMA0851 [imported] - Neisseria mening
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
C:Accession: G81930
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.;
Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajan
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22

```

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: G81930

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-177 <PAR>

A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84132.1; PID:g737956

A:Experimental source: serogroup A, strain 22491

C:Genetics:

A:Gene: pps; NMA0851

C:Superfamily: inorganic pyrophosphatase

C:Keywords: hydrolase

Query Match 47.4%; Score 45; DB 2; Length 177;

Best Local Similarity 43.8%; Pred. No. 6.4; 3; Indels 0; Gaps 0;

Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 2 DADQRIIKOLEYFNG 17

Db 126 DLPQOLIKOIEFHNH 141

RESULT 12

A56678

Yemanuclein-alpha - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C:Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 21-Jun-2002

A:Accession: A56678; S22146

R:Alt-Ahmed, O.; Bellon, B.; Capri, M.; Joblet, C.; Thomas-Delaage, M.

Mech. Dev. 37, 69-80, 1992

A:Title: The yemanuclein-alpha: a new *Drosophila* DNA binding protein specific for the oc

A:Reference number: A56678; MUID:92297435; PMID:1606021

A:Accession: A56678

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1002 <AIE>

A:Cross-references: GB:X63503; NID:g8837; PID:g8838

C:Genetics:

A:Gene: FlyBase:yem-alpha

A:Cross-references: FlyBase:FBgn0005596

A:Introns: 80/3; 154/3; 428/1 477/2; 557/2

C:Keywords: DNA binding; oocyte

Query Match 47.4%; Score 45; DB 2; Length 1002;

Best Local Similarity 52.9%; Pred. No. 42;

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 DDADQRIIKOLEYFNG 17

Db 105 DDDVARIKLEAKYGN 121

RESULT 13

F64526

adenine/cytosine DNA methyltransferase - *Helicobacter pylori* (strain 26695)

C:Species: *Helicobacter pylori*

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

A:Accession: F64526

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: F64526

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-823 <TOM>

A:Cross-references: GB:AE000527; GB:AE000511; NID:g2313131; PIDN:AD007124.1; PID:g231313

Query Match

Best Local Similarity 46.8%; Score 44.5; DB 2; Length 823;

Matches 8; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

Qy 2 DADQRIIKOLEYFNG 17

Db 65 ETEKILKQIEFYSKKFGN 83

RESULT 14

AD2410

hypothetical protein alr4836 [imported] - *Nostoc* sp. (strain PCC 7120)

C:Species: *Nostoc* sp.

A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C:Accession: AD2410

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iri

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tab

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacteri

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AD2410

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-158 <KUR>

A:Cross-references: GB:BA000019; PIDN:BA076535.1; PID:g17133973; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr4836

Query Match 46.3%; Score 44; DB 2; Length 158;

Best Local Similarity 50.0%; Pred. No. 8.2;

Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 5 QRIIKOLEYFNGI 18

Db 19 QRLIKSFQWTGNL 32

RESULT 15

AI3101

transcription regulator, *Lysr* family Atu437 [imported] - *Agrobacterium tumefaciens*

C:Species: *Agrobacterium tumefaciens*

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 02-Aug-2002

C:Accession: AI3101

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; M

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; M

Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AI3101

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-300 <KUR>

A:Cross-references: GB:AE008689; PIDN:AAL45231.1; PID:g17742913; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu437

A:Map position: linear chromosome

C:Superfamily: probable transcription regulator *lsyR*

Query Match 46.3%; Score 44; DB 2; Length 300;

Best Local Similarity 60.0%; Pred. No. 16;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 ADQRIIKOLEYFNG 17

Db 31 AVSRQIENLEYFNGS 45

Search completed: April 23, 2003, 13:34:41

Job time : 11.1124 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:27:28 ; Search time 10.3146 Seconds
(without alignments)
139.834 Million cell updates/sec

Title: US-09-836-073-17
Perfect score: 95
Sequence: 1 DDADQRIKOLEYFGNI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues
Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	100.0	18	9	US-09-836-073-17
2	55	57.9	18	9	US-09-836-073-15
3	50	52.6	16	9	US-09-836-073-19
4	48	50.5	397	9	US-09-738-626-5064
5	46	48.4	39	9	US-09-843-676-26
6	46	48.4	39	9	US-09-766-253-26
7	46	48.4	39	9	US-09-438-486-26
8	46	48.4	39	9	US-10-053-758-26
9	46	48.4	39	9	US-10-054-295-26
10	46	48.4	39	9	US-10-054-611-26
11	45	47.4	18	9	US-09-836-073-2
12	45	47.4	223	12	US-10-005-168-2
13	45	47.4	1076	9	US-10-116-949-6
14	45	47.4	1145	9	US-10-116-949-2
15	45	47.4	1145	9	US-10-116-949-4
16	44.5	46.8	823	9	US-09-895-913A-244
17	44	46.3	17	9	US-09-836-073-13
18	44	46.3	18	9	US-09-836-073-1
19	44	46.3	18	9	US-09-836-073-4

20	44	46.3	18	9	US-09-836-073-9	Sequence 9, Appli
21	44	46.3	18	9	US-09-836-073-14	Sequence 14, Appl
22	44	46.3	198	9	US-09-738-626-5036	Sequence 5036, Ap
23	44	46.3	460	9	US-10-102-806-695	Sequence 695, App
24	43	45.3	18	9	US-09-836-073-5	Sequence 5, Appli
25	43	45.3	38	9	US-09-843-676-25	Sequence 25, Appl
26	43	45.3	38	9	US-09-766-253-25	Sequence 25, Appl
27	43	45.3	38	9	US-09-438-486-25	Sequence 25, Appl
28	43	45.3	38	9	US-10-053-758-25	Sequence 25, Appl
29	43	45.3	38	9	US-10-054-295-25	Sequence 25, Appl
30	43	45.3	38	9	US-10-054-611-25	Sequence 25, Appl
31	41	43.2	185	10	US-09-815-837-13	Sequence 13, Appl
32	41	43.2	212	10	US-09-815-837-37	Sequence 37, Appl
33	41	43.2	361	10	US-09-815-242-10235	Sequence 10235, A
34	40	42.1	108	9	US-09-809-391-552	Sequence 552, App
35	40	42.1	186	9	US-09-809-391-394	Sequence 394, App
36	39.5	41.6	336	10	US-09-789-919-52	Sequence 52, Appl
37	39	41.1	18	9	US-09-836-073-3	Sequence 3, Appli
38	38	40.0	31	10	US-09-864-761-33555	Sequence 33555, A
39	38	40.0	461	10	US-09-764-864-1409	Sequence 1409, Ap
40	38	40.0	566	10	US-09-829-482-5	Sequence 5, Appli
41	38	40.0	576	10	US-09-891-216-13	Sequence 13, Appl
42	38	40.0	684	10	US-09-891-216-14	Sequence 14, Appl
43	38	40.0	732	9	US-10-008-355-6	Sequence 6, Appli
44	38	40.0	934	10	US-09-891-216-12	Sequence 12, Appl
45	38	40.0	934	10	US-09-891-216-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-09-836-073-17
; Sequence 17, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220003054822
; CURRENT APPLICATION NUMBER: US/09/836, 073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316, 630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 18
; TYPE: PRT
; ORGANISM: C. elegans
US-09-836-073-17
Query Match 100.0%; Score 95; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DDADQRIKOLEYFGNI 18
Db 1 DDADQRIKOLEYFGNI 18
RESULT 2
US-09-836-073-15
; Sequence 15, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836, 073

SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-843-676-26

Query Match 48.4%; Score 46; DB 9; Length 39;
Best Local Similarity 63.6%; Pred. No. 0.85;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 IIKQLEYFGN 17
:::|||||
Db 1 ILRQVEYFGD 11

RESULT 6

US-09-766-253-26

; Sequence 26, Application US/09766253
; Publication No. US20020187471A1

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
; TITLE OF INVENTION: No. US20020187471A1 Telomerase

; NUMBER OF SEQUENCES: 171

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/766,253
; FILING DATE: 19-Jan-2001
; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/846,017
; FILING DATE: 1997-04-25
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002920US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 26:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid

; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 26:

US-09-766-253-26

Query Match 48.4%; Score 46; DB 9; Length 39;
Best Local Similarity 63.6%; Pred. No. 0.85;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 IIKQLEYFGN 17
:::|||||
Db 1 ILRQVEYFGD 11

RESULT 7

US-09-438-486-26

; Sequence 26, Application US/09438486
; Publication No. US20030009019A1

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. US20030009019A1 Telomerase

; NUMBER OF SEQUENCES: 223

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/438,486
; FILING DATE: 12-NOV-1999
; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002931US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 26:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid

; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant

; MOLECULE TYPE: peptide

US-09-438-486-26

Query Match 48.4%; Score 46; DB 9; Length 39;

Best Local Similarity 63.6%; Pred. No. 0.85;

Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 IIKQLEYFGN 17

:::|||||

Db 1 ILRQVEYFGD 11

RESULT 8

US-10-053-758-26

```
; Sequence 26, Application US/10053758
; Publication No. US20030032075A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
; TITLE OF INVENTION: NO. US20030032075A1 Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/053,758
; FILING DATE: 18-Jan-2002
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030032075A1 Relevant
; TOPOLOGY: No. US20030032075A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-10-053-758-26

Query Match 48.4%; Score 46; DB 9; Length 39;
Best Local Similarity 63.6%; Pred. No. 0.85;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 IIKQLEYVFGN 17
Db 1 ILRQVEYVFGD 11

RESULT 9
US-10-054-295-26
; Sequence 26, Application US/10054295
; Publication No. US20030044953A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
; TITLE OF INVENTION: NO. US20030044953A1 Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,295
; FILING DATE: 18-Jan-2002
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030044953A1 Relevant
; TOPOLOGY: No. US20030044953A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-10-054-295-26

Query Match 48.4%; Score 46; DB 9; Length 39;
Best Local Similarity 63.6%; Pred. No. 0.85;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 IIKQLEYVFGN 17
Db 1 ILRQVEYVFGD 11

RESULT 10
US-10-054-611-26
; Sequence 26, Application US/10054611
; Publication No. US20030059787A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
; TITLE OF INVENTION: NO. US20030059787A1 Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
```

STREET: Two Embarcadero Center, 8th Floor
 CITY: San Francisco
 STATE: California
 COUNTRY: United States of America
 ZIP: 94111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/054,611
 FILING DATE: 18-Jan-2002
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/854,050
 FILING DATE: <Unknown>
 APPLICATION NUMBER: US 08/846,017
 FILING DATE: 25-APR-1997
 APPLICATION NUMBER: US 08/844,419
 FILING DATE: 18-APR-1997
 APPLICATION NUMBER: US 08/724,643
 FILING DATE: 01-OCT-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Apple, Randolph T.
 REGISTRATION NUMBER: 36,429
 REFERENCE/DOCKET NUMBER: 015389-002930US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 26:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 39 amino acids
 TYPE: amino acid
 STRANDEDNESS: No. US20030059787A1 Relevant
 TOPOLOGY: No. US20030059787A1 Relevant
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 26:
 US-10-054-611-26

Query Match 48.4%; Score 46; DB 9; Length 39;
 Best Local Similarity 63.6%; Pred. No. 0.85;
 Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 7 IIKOLEYYFGN 17
 Db 1 ILRQVEYYFGD 11

RESULT 11
 US-09-836-073-2
 ; Sequence 2, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Balaya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 22002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-2

Query Match 47.4%; Score 45; DB 9; Length 18;
 Best Local Similarity 50.0%; Pred. No. 0.55;

Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 Qy 4 DQRIIKOLEYYFGN 17
 Db 4 EAQICQOIEYFGD 17

RESULT 12
 US-10-005-168-2
 ; Sequence 2, Application US/10005168
 ; Patent No. US20020133840A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Brennan, Thomas J.
 ; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING DISRUPTIONS
 ; TITLE OF INVENTION: IN GENES ENCODING A2D2 CALCIUM ION CHANNEL SUBUNIT PROTEINS
 ; FILE REFERENCE: R-10
 ; CURRENT APPLICATION NUMBER: US/10/005,168
 ; CURRENT FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: US 60/299,668
 ; PRIOR FILING DATE: 2001-06-19
 ; PRIOR APPLICATION NUMBER: US 60/282,685
 ; PRIOR FILING DATE: 2001-04-09
 ; PRIOR APPLICATION NUMBER: US 60/254,802
 ; PRIOR FILING DATE: 2000-12-11
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 223
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-10-005-168-2

Query Match 47.4%; Score 45; DB 12; Length 223;
 Best Local Similarity 50.0%; Pred. No. 7.5;
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 7 IIKOLEYYFGNI 18
 Db 86 VMKQTQYFGSV 97

RESULT 13
 US-10-116-949-6
 ; Sequence 6, Application US/10116949
 ; Publication No. US20030044911A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lerman, Michael I.
 ; APPLICANT: Minna, John D.
 ; APPLICANT: Latif, Farida
 ; APPLICANT: Wei, Ming-Hui
 ; APPLICANT: Sekido, Yoshitaka
 ; APPLICANT: Gao, Boning
 ; APPLICANT: Duh, Fuh-Mei
 ; TITLE OF INVENTION: Calcium Channel Compositions and Methods of Use Thereof
 ; FILE REFERENCE: NIH-05043
 ; CURRENT APPLICATION NUMBER: US/10/116,949
 ; CURRENT FILING DATE: 2002-04-05
 ; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-12-22
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-22
 ; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1998-12-30
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-30
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 6
 ; LENGTH: 1076
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-116-949-6

Query Match 47.4%; Score 45; DB 9; Length 1076;
 Best Local Similarity 50.0%; Pred. No. 38;
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 7 IIKQLEYFGNI 18
 :|| :||||:
 Db 938 VMKQTYFGSV 949

RESULT 14

US-10-116-949-2
 ; Sequence 2, Application US/10116949
 ; Publication NO. US2003004911A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lerman, Michael I.
 ; APPLICANT: Minna, John D.
 ; APPLICANT: Latif, Farida
 ; APPLICANT: Wei, Ming-Hui
 ; APPLICANT: Sekido, Yoshitaka
 ; APPLICANT: Gao, Boning
 ; APPLICANT: Duh, Fuh-Mei
 ; TITLE OF INVENTION: Calcium Channel Compositions and Methods of Use Thereof
 ; FILE REFERENCE: NIH-05043
 ; CURRENT APPLICATION NUMBER: US/10/116,949
 ; CURRENT FILING DATE: 2002-04-05
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/470,443
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-22
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/114,359
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-30
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 1145
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-116-949-2

Query Match 47.4%; Score 45; DB 9; Length 1145;
 Best Local Similarity 50.0%; Pred. No. 41;
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 7 IIKQLEYFGNI 18
 :|| :||||:
 Db 1007 VMKQTYFGSV 1018

RESULT 15

US-10-116-949-4
 ; Sequence 4, Application US/10116949
 ; Publication NO. US2003004911A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lerman, Michael I.
 ; APPLICANT: Minna, John D.
 ; APPLICANT: Latif, Farida
 ; APPLICANT: Wei, Ming-Hui
 ; APPLICANT: Sekido, Yoshitaka
 ; APPLICANT: Gao, Boning
 ; APPLICANT: Duh, Fuh-Mei
 ; TITLE OF INVENTION: Calcium Channel Compositions and Methods of Use Thereof
 ; FILE REFERENCE: NIH-05043
 ; CURRENT APPLICATION NUMBER: US/10/116,949
 ; CURRENT FILING DATE: 2002-04-05
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/470,443
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-22
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/114,359
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-30
 ; NUMBER OF SEQ ID NOS: 114
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 1145
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-116-949-4

Query Match 47.4%; Score 45; DB 9; Length 1145;
 Best Local Similarity 50.0%; Pred. No. 41;
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 7 IIKQLEYFGNI 18
 :|| :||||:
 Db 1007 VMKQTYFGSV 1018

Search completed: April 23, 2003, 13:38:22
 Job time : 11.3146 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:24 ; Search time 11.2247 Seconds
(without alignments)
47.183 Million cell updates/sec

Title: US-09-836-073-17
Perfect score: 95
Sequence: 1 DDADQRIIKQLEYFGNI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	50.5	204	2	US-08-516-801-3
2	48	50.5	204	4	US-08-248-355-3
3	48	50.5	204	5	PCT-US95-06683-3
4	48	50.5	385	2	US-08-516-801-2
5	48	50.5	385	4	US-08-248-355-2
6	48	50.5	385	5	PCT-US95-06683-2
7	47	49.5	108	4	US-09-187-859-41
8	46	48.4	39	3	US-08-851-843A-26
9	46	48.4	39	4	US-08-974-549A-216
10	46	48.4	39	4	US-08-854-050-26
11	46	48.4	39	4	US-09-430-323-26
12	45	47.4	1076	4	US-09-470-443-6
13	45	47.4	1145	4	US-09-470-443-2
14	45	47.4	1145	4	US-09-470-443-4
15	44	46.3	18	4	US-09-316-630-3
16	44	46.3	18	4	US-09-316-630-4
17	43	45.3	38	3	US-08-851-843A-25
18	43	45.3	38	4	US-08-974-549A-215
19	43	45.3	38	4	US-08-854-050-25
20	43	45.3	38	4	US-09-430-323-25
21	41	43.2	108	4	US-09-187-859-43
22	41	43.2	775	3	US-08-966-388-4
23	41	43.2	775	3	US-09-188-403-4
24	41	43.2	775	4	US-09-188-404-4
25	41	43.2	775	4	US-09-281-259-4
26	40	42.1	108	4	US-09-149-476-552
27	40	42.1	185	3	US-08-984-295-1

28 40 42.1 186 4 US-09-149-476-394
29 38 40.0 38 4 US-08-974-549A-214
30 38 40.0 156 2 US-09-098-900-4
31 38 40.0 156 2 US-08-806-877-4
32 38 40.0 367 3 US-09-226-741-3
33 38 40.0 367 4 US-09-595-514-3
34 38 40.0 438 1 US-08-439-131A-2
35 38 40.0 438 1 US-08-440-674-5
36 38 40.0 490 3 US-09-226-741-1
37 38 40.0 490 4 US-09-595-514-1
38 38 40.0 566 4 US-09-073-297-5
39 38 40.0 1319 4 US-09-462-136-4
40 37 38.9 83 3 US-08-851-843A-9
41 37 38.9 83 4 US-08-974-549A-191
42 37 38.9 83 4 US-08-854-050-9
43 37 38.9 83 4 US-09-430-323-9
44 37 38.9 85 3 US-08-851-843A-11
45 37 38.9 85 4 US-08-974-549A-193

ALIGNMENTS

RESULT 1
US-08-516-801-3
; Sequence 3, Application US/08516801
; Patent No. 5872213
; GENERAL INFORMATION:

; APPLICANT: Goff, Stephen P.
; TITLE OF INVENTION: A cdna Clone of the Human Ini-1 Gene Encoding a
; TITLE OF INVENTION: Protein That Binds to the HIV-1 Integrase; and Its Use in
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/516.801
; FILING DATE: 18-August-1995

; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: John P. White, Esq.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 44301-2/JPW/AGL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 204 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FRAGMENT TYPE: N-terminal
; US-08-516-801-3

Query Match 50.5%; Score 48; DB 2; Length 204;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIIKQLEYFGNI 18

Db 77 DQSDQRIKLEIYFNGI 94
| : ||| : | : |||

RESULT 2

US-08-248-355-3
; Sequence 3, Application US/08248355
; Patent No. 622024
; GENERAL INFORMATION:
; APPLICANT: Goff, Stephen P.
; TITLE OF INVENTION: A cDNA Clone of the Human Ini-1 Gene Encoding
; TITLE OF INVENTION: a Protein That Binds to the HIV-1 Integrase;
; TITLE OF INVENTION: and Its Use in Antiviral Therapy
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,355
; FILING DATE: 24-May-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq., John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 44301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 204 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FRAGMENT TYPE: N-terminal
US-08-248-355-3

Query Match 50.5%; Score 48; DB 4; Length 204;
Best Local Similarity 50.0%; Pred. NO. 1.2;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIKLEIYFNGI 18
| : ||| : | : |||
Db 77 DQSDQRIKLEIYFNGI 94

RESULT 3

PCT-US95-06683-3
; Sequence 3, Application PC/TUS9506683
; GENERAL INFORMATION:
; APPLICANT: Goff, Stephen P.
; APPLICANT: Kalpana, Ganjam V.
; TITLE OF INVENTION: A cDNA Clone of the Human Ini-1 Gene Encoding
; TITLE OF INVENTION: a Protein That Binds to the HIV-1 Integrase;
; TITLE OF INVENTION: and Its Use in Antiviral Therapy
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York

; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06683
; FILING DATE: 24-May-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/248,355
; FILING DATE: 24-May-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq., John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 44301-A-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 204 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FRAGMENT TYPE: N-terminal
PCT-US95-06683-3

Query Match 50.5%; Score 48; DB 5; Length 204;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIKLEIYFNGI 18
| : ||| : | : |||
Db 77 DQSDQRIKLEIYFNGI 94

RESULT 4

US-08-516-801-2
; Sequence 2, Application US/08516801
; Patent No. 5872213
; GENERAL INFORMATION:
; APPLICANT: Goff, Stephen P.
; APPLICANT: Kalpana, Ganjam V.
; TITLE OF INVENTION: A cDNA Clone of the Human Ini-1 Gene Encoding a
; TITLE OF INVENTION: Protein That Binds to the HIV-1 Integrase; and Its Use in

Query Match 50.5%; Score 48; DB 4; Length 204;
Best Local Similarity 50.0%; Pred. NO. 1.2;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIKLEIYFNGI 18
| : ||| : | : |||
Db 77 DQSDQRIKLEIYFNGI 94

RESULT 3

PCT-US95-06683-3
; Sequence 3, Application PC/TUS9506683
; GENERAL INFORMATION:
; APPLICANT: Goff, Stephen P.
; APPLICANT: Kalpana, Ganjam V.
; TITLE OF INVENTION: A cDNA Clone of the Human Ini-1 Gene Encoding
; TITLE OF INVENTION: a Protein That Binds to the HIV-1 Integrase;
; TITLE OF INVENTION: and Its Use in Antiviral Therapy
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York

REFERENCE/DOCKET NUMBER: 44301-Z/JPW/AGL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-516-801-2

Query Match 50.5%; Score 48; DB 2; Length 385;
Best Local Similarity 50.0%; Pred. No. 2.5;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIIKOLEYFGNI 18
| :|||:| :|||
Db 256 DQSDQVRVILKNIHVGN 273

RESULT 5

US-08-248-355-2
Sequence 2, Application US/08248355
Patent No. 6222024
GENERAL INFORMATION:

APPLICANT: Goff, Stephen P.
APPLICANT: Kalpana, Ganjam V.
TITLE OF INVENTION: A cDNA Clone of the Human Ini-1 Gene Encoding
TITLE OF INVENTION: a Protein That Binds to the HIV-1 Integrase;
TITLE OF INVENTION: and Its Use in Antiviral Therapy
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/248,355
FILING DATE: 24-May-1994
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: White Esq., John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 44301
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-248-355-2

Query Match 50.5%; Score 48; DB 4; Length 385;
Best Local Similarity 50.0%; Pred. No. 2.5;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIIKOLEYFGNI 18
| :|||:| :|||
Db 256 DQSDQVRVILKNIHVGN 273

RESULT 6

US-08-248-355-2
Sequence 2, Application US/08248355
Patent No. 6222024
GENERAL INFORMATION:

APPLICANT: Goff, Stephen P.
APPLICANT: Kalpana, Ganjam V.
TITLE OF INVENTION: A cDNA Clone of the Human Ini-1 Gene Encoding
TITLE OF INVENTION: a Protein That Binds to the HIV-1 Integrase;
TITLE OF INVENTION: and Its Use in Antiviral Therapy
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06683
FILING DATE: 24-May-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/248,355
FILING DATE: 24-May-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White Esq., John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 44301-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-06683-2

Query Match 50.5%; Score 48; DB 5; Length 385;
Best Local Similarity 50.0%; Pred. No. 2.5;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIIKOLEYFGNI 18
| :|||:| :|||
Db 256 DQSDQVRVILKNIHVGN 273

PCT-US95-06683-2
Sequence 2, Application PC/TUS9506683
GENERAL INFORMATION:

APPLICANT: Goff, Stephen P.
APPLICANT: Kalpana, Ganjam V.
TITLE OF INVENTION: A cDNA Clone of the Human Ini-1 Gene Encoding
TITLE OF INVENTION: a Protein That Binds to the HIV-1 Integrase;
TITLE OF INVENTION: and Its Use in Antiviral Therapy
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06683
FILING DATE: 24-May-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/248,355
FILING DATE: 24-May-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White Esq., John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 44301-A-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-06683-2

QY 1 DDADQRIIKOLEYFGNI 18
| :|||:| :|||
Db 256 DQSDQVRVILKNIHVGN 273

RESULT 7

US-09-187-859-41
Sequence 41, Application US/09187859A
Patent No. 6358920
GENERAL INFORMATION:

APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
FILE REFERENCE: 100086.407C1
CURRENT APPLICATION NUMBER: US/09/187,859A
CURRENT FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 4052
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 41
LENGTH: 108
TYPE: PRT
ORGANISM: Mus musculus
US-09-187-859-41

Query Match 50.5%; Score 48; DB 5; Length 385;
Best Local Similarity 50.0%; Pred. No. 2.5;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIIKOLEYFGNI 18
| :|||:| :|||
Db 256 DQSDQVRVILKNIHVGN 273

Query Match 49.5%; Score 47; DB 4; Length 108;
Best Local Similarity 47.1%; Pred. No. 0.88;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 DADQRIIKOLEYVFGNI 18
|||:|:|:|:|:|:
Db 26 DADGINKETLYFESNL 42

RESULT 8
US-08-851-843A-26
; Sequence 26, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-851-843A-26

Query Match 48.4%; Score 46; DB 3; Length 39;
Best Local Similarity 63.8%; Pred. No. 0.41;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 7 IIKOLEYVFGN 17
|:|:|:|:|:|:

Db 1 ILRQVEYVFGD 11

RESULT 9
US-08-974-549A-216
; Sequence 216, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 216:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids

;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-974-549A-216

Query Match 48.4%; Score 46; DB 4; Length 39;
Best Local Similarity 63.6%; Pred. No. 0.41;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 IIKOLEYYFGN 17
Db 1 ILRQVEYYFGD 11

RESULT 10
US-08-854-050-26
; Sequence 26, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: NO. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-0029300S
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-08-430-323-26

Query Match 48.4%; Score 46; DB 4; Length 39;
Best Local Similarity 63.6%; Pred. No. 0.41;

;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
US-08-854-050-26

Query Match 48.4%; Score 46; DB 4; Length 39;
Best Local Similarity 63.6%; Pred. No. 0.41;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 IIKOLEYYFGN 17
Db 1 ILRQVEYYFGD 11

RESULT 11
US-09-430-323-26
; Sequence 26, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: NO. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-0029300S
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-430-323-26

Query Match 48.4%; Score 46; DB 4; Length 39;
Best Local Similarity 63.6%; Pred. No. 0.41;

Matches: 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 IIKOLEYFNG 17
I::I::I::I::I::
Db 1 ILQVEYFGD 11

RESULT 12
US-09-470-443-6
; Sequence 6, Application US/09470443
; Patent No. 6441156
; GENERAL INFORMATION:
; APPLICANT: Lerman, Michael I.
; APPLICANT: Minna, John D.
; APPLICANT: Latif, Farida
; APPLICANT: Wei, Ming-Hui
; APPLICANT: Sekido, Yoshitaka
; APPLICANT: Gao, Boning
; APPLICANT: Duh, Fuh-Mei
; TITLE OF INVENTION: Calcium Channel Compositions and Methods of Use Thereof
; FILE REFERENCE: NIH-05043
; CURRENT APPLICATION NUMBER: US/09/470,443
; CURRENT FILING DATE: 1999-12-22
; EARLIER APPLICATION NUMBER: 60/114,359
; EARLIER FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1076
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-470-443-6

Query Match 47.4%; Score 45; DB 4; Length 1076;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 7 IIKOLEYFNG 18
I::I::I::I::I::
Db 938 VMKOTQYFGSV 949

RESULT 13
US-09-470-443-2
; Sequence 2, Application US/09470443
; Patent No. 6441156
; GENERAL INFORMATION:
; APPLICANT: Lerman, Michael I.
; APPLICANT: Minna, John D.
; APPLICANT: Latif, Farida
; APPLICANT: Wei, Ming-Hui
; APPLICANT: Sekido, Yoshitaka
; APPLICANT: Gao, Boning
; APPLICANT: Duh, Fuh-Mei
; TITLE OF INVENTION: Calcium Channel Compositions and Methods of Use Thereof
; FILE REFERENCE: NIH-05043
; CURRENT APPLICATION NUMBER: US/09/470,443
; CURRENT FILING DATE: 1999-12-22
; EARLIER APPLICATION NUMBER: 60/114,359
; EARLIER FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-470-443-2

Query Match 47.4%; Score 45; DB 4; Length 1145;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 7 IIKOLEYFNG 18
I::I::I::I::I::

Db 1007 VMKOTQYFGSV 1018
I::I::I::I::I::

RESULT 14
US-09-470-443-4
; Sequence 4, Application US/09470443
; Patent No. 6441156
; GENERAL INFORMATION:
; APPLICANT: Lerman, Michael I.
; APPLICANT: Minna, John D.
; APPLICANT: Latif, Farida
; APPLICANT: Wei, Ming-Hui
; APPLICANT: Sekido, Yoshitaka
; APPLICANT: Gao, Boning
; APPLICANT: Duh, Fuh-Mei
; TITLE OF INVENTION: Calcium Channel Compositions and Methods of Use Thereof
; FILE REFERENCE: NIH-05043
; CURRENT APPLICATION NUMBER: US/09/470,443
; CURRENT FILING DATE: 1999-12-22
; EARLIER APPLICATION NUMBER: 60/114,359
; EARLIER FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-470-443-4

Query Match 47.4%; Score 45; DB 4; Length 1145;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 7 IIKOLEYFNG 18
I::I::I::I::I::
Db 1007 VMKOTQYFGSV 1018

RESULT 15
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match 46.3%; Score 44; DB 4; Length 18;
Best Local Similarity 50.0%; Pred. No. 0.37;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 DORIIKOLEYFNG 17
I::I::I::I::I::

Db 4 EAKICHQIEYFGD 17

Search completed: April 23, 2003, 13:36:34
Job time : 12.2247 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:23 ; Search time 28.5169 Seconds
(without alignments)
84.109 Million cell updates/sec

Title: US-09-836-073-17

Perfect score: 95

Sequence: 1 DDADQRIKQLEYFNGNI 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	56.8	913	21	Arabidopsis thalia
2	54	56.8	923	21	Arabidopsis thalia
3	54	56.8	993	21	Arabidopsis thalia
4	50	52.6	390	22	Arabidopsis thalia
5	48	50.5	376	22	Drosophila melanog
6	48	50.5	384	21	Human protein sequ
7	48	50.5	385	17	Human Ini-1. Homo
8	48	50.5	394	22	Human integrase in
9	48	50.5	397	22	Human protein sequ
10	48	50.5	650	22	C glutamicum prote
					Novel human diagno

11	47	49.5	108	21	AAV64613	Nonclassical cadhe
12	45	47.4	198	22	RAM25632	Human protein sequ
13	45	47.4	1002	22	ABB58137	Drosophila melanog
14	45	47.4	1062	22	AAU01014	Human secreted sol
15	45	47.4	1062	22	ABB62238	Human calcium chan
16	45	47.4	1082	22	AAU01015	Human secreted sol
17	45	47.4	1082	22	ABB62239	Human calcium chan
18	45	47.4	1109	22	AAU01016	Human secreted sol
19	45	47.4	1109	22	ABB62240	Human calcium chan
20	45	47.4	1115	22	AAU01025	Human wild type al
21	45	47.4	1115	22	ABB62249	Human calcium chan
22	45	47.4	1145	21	AAV92322	Human alpha-2-delt
23	45	47.4	1145	22	AAU01023	Human secreted sol
24	45	47.4	1145	22	ABB62247	Human calcium chan
25	44.5	46.8	823	19	AAV98507	H. pylori GHP0 122
26	44	46.3	18	21	AAV52200	Human la autoantig
27	44	46.3	92	21	AAU01351	Human secreted pro
28	44	46.3	198	22	AAU01282	C glutamicum prote
29	44	46.3	408	17	AAU03716	Human autoantigen
30	44	46.3	439	22	ABB08417	Novel human diagno
31	44	46.3	460	21	AAU58987	Breast and ovarian
32	44	46.3	460	23	ABP41511	Human ovarian anti
33	43	45.3	538	22	ABB65701	Drosophila melanog
34	42	44.2	79	20	AAV22534	Drosophila HBP(p9)
35	42	44.2	136	22	AAU93351	Human polypeptide,
36	42	44.2	363	22	ABB64805	Drosophila melanog
37	42	44.2	365	22	ABB66075	Drosophila melanog
38	42	44.2	514	22	ABB52689	Escherichia coli p
39	41.5	43.7	598	13	AAU20470	Soluble ysc alpha.
40	41.5	43.7	729	11	AAU05457	KE1 amino acid se
41	41.5	43.7	827	22	AAU62909	Amino acid sequenc
42	41	43.2	21	14	AAU43394	La/SSB epitope 17.
43	41	43.2	108	21	AAV64615	Nonclassical cadhe
44	41	43.2	156	22	ABB29928	Novel human diagno
45	41	43.2	185	22	ABB56427	Human single chain

ALIGNMENTS

RESULT 1
AAG47714 standard; Protein; 913 AA.

AC AAG47714;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 60168.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

PR	21-APR-1999;	99US-0130449.
PR	23-APR-1999;	99US-0130510.
PR	23-APR-1999;	99US-0130891.
PR	28-APR-1999;	99US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	07-MAY-1999;	99US-0132487.
PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	17-JUN-1999;	99US-0139453.
PR	18-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	24-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	02-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	28-JUL-1999;	99US-0145919.
PR	02-AUG-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	30-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	07-OCT-1999;	99US-0157865.
PR	08-OCT-1999;	99US-0158029.
PR	12-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.


```
PR 22-JUL-1999; 99US-0145087;
PR 22-JUL-1999; 99US-0145089;
PR 22-JUL-1999; 99US-0145192;
PR 23-JUL-1999; 99US-0145145;
PR 23-JUL-1999; 99US-0145218;
PR 23-JUL-1999; 99US-0145224;
PR 26-JUL-1999; 99US-0145276;
PR 27-JUL-1999; 99US-0145913;
PR 27-JUL-1999; 99US-0145918;
PR 27-JUL-1999; 99US-0145919;
PR 28-JUL-1999; 99US-0145919;
PR 02-AUG-1999; 99US-0145951;
PR 02-AUG-1999; 99US-0146386;
PR 02-AUG-1999; 99US-0146388;
PR 02-AUG-1999; 99US-0146389;
PR 03-AUG-1999; 99US-0147038;
PR 04-AUG-1999; 99US-0147204;
PR 04-AUG-1999; 99US-0147302;
PR 05-AUG-1999; 99US-0147192;
PR 05-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149175;
PR 18-AUG-1999; 99US-0149426;
PR 20-AUG-1999; 99US-0149722;
PR 20-AUG-1999; 99US-0149723;
PR 20-AUG-1999; 99US-0149929;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151085;
PR 27-AUG-1999; 99US-0151086;
PR 27-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151930;
PR 07-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155119;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;
PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159329;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 14-OCT-1999; 99US-0159638;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0160741;
PR 21-OCT-1999; 99US-0160767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 21-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161922;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;
PR 29-OCT-1999; 99US-0162143;
Query Match 56.8%; Score 54; DB 21; Length 923;
Best Local Similarity 44.4%; Pred. No. 2.5;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
Qy 1 DDADQRIIKQLEYFNGI 18
Db 121 EDSIQKIVNQVEYFSDL 138
RESULT 3
AAG47712
ID AAG47712 standard; Protein; 993 AA.
XX
AC AAG47712;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 60166.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825;
PR 05-MAR-1999; 99US-0123180;
PR 09-MAR-1999; 99US-0123548;
PR 23-MAR-1999; 99US-0125788;
PR 25-MAR-1999; 99US-0126264;
PR 29-MAR-1999; 99US-0126785;
PR 01-APR-1999; 99US-0127462;
PR 06-APR-1999; 99US-0128234;
PR 08-APR-1999; 99US-0128714;
PR 16-APR-1999; 99US-0129845;
PR 19-APR-1999; 99US-0130077;
PR 21-APR-1999; 99US-0130449;
PR 23-APR-1999; 99US-0130510;
PR 23-APR-1999; 99US-0130891;
PR 28-APR-1999; 99US-0131449;
PR 30-APR-1999; 99US-0132048;
PR 30-APR-1999; 99US-0132407;
PR 04-MAY-1999; 99US-0132484;
PR 05-MAY-1999; 99US-0132485;
PR 06-MAY-1999; 99US-0132486;
PR 07-MAY-1999; 99US-0132487;
PR 11-MAY-1999; 99US-0132863;
PR 14-MAY-1999; 99US-0134256;
PR 14-MAY-1999; 99US-0134218;
PR 14-MAY-1999; 99US-0134219;
```

PR 14-MAY-1999;	99US-0134221.	PR 27-JUL-1999;	99US-0145913.
PR 14-MAY-1999;	99US-0134370.	PR 27-JUL-1999;	99US-0145918.
PR 18-MAY-1999;	99US-0134768.	PR 27-JUL-1999;	99US-0145919.
PR 19-MAY-1999;	99US-0134941.	PR 27-JUL-1999;	99US-0145951.
PR 20-MAY-1999;	99US-0135124.	PR 02-AUG-1999;	99US-0146386.
PR 21-MAY-1999;	99US-0135353.	PR 02-AUG-1999;	99US-0146388.
PR 24-MAY-1999;	99US-0135629.	PR 02-AUG-1999;	99US-0146389.
PR 25-MAY-1999;	99US-0136021.	PR 03-AUG-1999;	99US-0147038.
PR 27-MAY-1999;	99US-0136392.	PR 04-AUG-1999;	99US-0147204.
PR 28-MAY-1999;	99US-0136782.	PR 04-AUG-1999;	99US-0147302.
PR 01-JUN-1999;	99US-0137222.	PR 05-AUG-1999;	99US-0147302.
PR 03-JUN-1999;	99US-0137528.	PR 05-AUG-1999;	99US-0147260.
PR 04-JUN-1999;	99US-0137502.	PR 06-AUG-1999;	99US-0147303.
PR 07-JUN-1999;	99US-0137724.	PR 06-AUG-1999;	99US-0147416.
PR 08-JUN-1999;	99US-0138094.	PR 09-AUG-1999;	99US-0147493.
PR 10-JUN-1999;	99US-0138540.	PR 09-AUG-1999;	99US-0147935.
PR 10-JUN-1999;	99US-0138847.	PR 10-AUG-1999;	99US-0148171.
PR 14-JUN-1999;	99US-0139119.	PR 11-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139452.	PR 12-AUG-1999;	99US-0148341.
PR 16-JUN-1999;	99US-0139453.	PR 13-AUG-1999;	99US-0148565.
PR 17-JUN-1999;	99US-0139492.	PR 13-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139454.	PR 16-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0139899.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140355.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140354.	PR 01-SEP-1999;	99US-0151930.
PR 24-JUN-1999;	99US-0140695.	PR 07-SEP-1999;	99US-0152363.
PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 29-JUN-1999;	99US-0140991.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154018.
PR 01-JUL-1999;	99US-0141842.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154779.
PR 02-JUL-1999;	99US-0142055.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 19-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0158293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.

Matches	8;	Conservative	5;	Mismatches	1;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	----	--------	----	------	----

SQ sequence 3/6 AA;

Query Match	50.5%;	Score 48;	DB 22;	Length 376;
Best Local Similarity	50.08;	Pred. No. 9.1;		

Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIKOLEYFNGI 18
|:||||:|:|
Db 247 DQSDQRIKLNTHVNGI 264

RESULT 6

AA92337
ID AAY92337 standard; Protein; 384 AA.

XX AC AAY92337;

DT 21-AUG-2000 (first entry)
DE Human Ini-1.

XX NLK1 Interacting Protein; Ini-1; protein complex; cytostatic; antiviral;
KW neuroprotective; cardiant.
XX Homo sapiens.

OS WO20002048-A2.

XX 13-APR-2000.

XX 06-OCT-1999; 99WO-US23314.

XX 06-OCT-1998; 98US-0167206.

XX (CURA-) CURAGEN CORP.

XX Nandabalan K, Schulz VP, Yang M;

XX WPI; 2000-303742/26.

XX N-PSDB; AAA09309.

XX New complex of a NLK1 protein and a NLK1 protein-interacting protein,
PT useful for treating cancer, hyperproliferative disorder,
PT neurodegenerative disorder, cardiomyopathies, viral infections and
PT metabolic disorders

PS Example 1; Page 162-163; 172pp; English.

XX AAY92331-37 were isolated in a modified yeast two hybrid system using
CC NLK1 protein as "bait". These are known sequences which are NLK1
CC interacting proteins. The invention concerns purified complexes of a
CC NLK1 protein and a NLK1 protein-interacting protein, where the
CC interacting protein is chosen from TrkA, protein phosphatase lalpa,
CC 14-3-3epsilon, alpha-tropomyosin, vimentin, p0071, Ini-1, Ip-1, Ip-2,
CC Ip-3, Ip-4, or Ip-5. NLK1 (also referred to as Nek2) is a human
CC homologue of the Aspergillus nidulans mitotic regulator, NIMA kinase.
CC NLK1 is a serine/threonine-specific kinase and is thought to play a key
CC role in cell-cycle events leading to the onset of mitosis. The
CC complexes, their derivatives and NLK1 or NLK1-IP protein and DNA
CC sequences, etc. are useful for treating or preventing a disease or
CC disorder involving aberrant levels of the complex or protein. Such
CC disorders include cancer, hyperproliferative disorders,
CC neurodegenerative disorders, cardiomyopathies, viral infections and
CC metabolic disorders.

XX Sequence 384 AA;

Query Match 50.5%; Score 48; DB 21; Length 384;
Best Local Similarity 50.0%; Pred. No. 9.3;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIKOLEYFNGI 18
|:||||:|:|
Db 256 DQSDQRIKLNTHVNGI 273

RESULT 7

AAW00081
ID AAW00081 standard; Protein; 385 AA.

XX AC AAW00081;

XX 13-NOV-1996 (first entry)
DT Human Integrase Interactor 1.

XX Integrase Interactor 1; ini-1 gene; antiviral; retrovirus; HIV-1;

XX human immunodeficiency virus type 1; therapy.

XX Homo sapiens.

XX WO9531904-A1.

XX 30-NOV-1995.

XX 24-MAY-1995; 95WO-US06683.

XX 24-MAY-1994; 94US-0248355.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Goff SP, Kalpana GV;

XX WPI; 1996-020286/02.

XX N-PSDB; AAT34357.

XX New isolated integrase Interactor-1 gene - used to develop assays
PT for anti-viral agents capable of interfering with integrase
PT Interactor protein binding
XX Example; Page 50-51; 73pp; English.

XX Human integrase Interactor 1 protein (AAW00081) was identified as
CC the product of the ini-1 gene (AAT34357) isolated from an HL60
CC cDNA library. The ini-1 protein may serve as an internal receptor
CC for retrovirus integrase enzymes, and may be responsible for
CC targeting integration to active regions of the chromosome. The
CC ini-1 protein can be expressed in prokaryotic or eukaryotic host
CC cells. It is useful for screening cpds. that disrupt retrovirus
CC (esp. HIV-1) integrase-1ni-1 protein interaction and thereby disrupt
CC the retrovirus life cycle.

SQ Sequence 385 AA;

Query Match 50.5%; Score 48; DB 17; Length 385;
Best Local Similarity 50.0%; Pred. No. 9.3;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIKOLEYFNGI 18
|:||||:|:|
Db 256 DQSDQRIKLNTHVNGI 273

RESULT 8

AA95622

ID AAB95622 standard; Protein; 394 AA.

XX AC AAB95622;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:18340.
DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.
 XX PR 29-JUL-1999; 99JP-0248036.
 XX PR 27-AUG-1999; 99JP-0300253.
 XX PR 11-JAN-2000; 2000JP-0118776.
 XX PR 02-MAY-2000; 2000JP-0183767.
 XX PR 09-JUN-2000; 2000JP-0241899.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX DR WPI: 2001-318749/34.
 XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
 XX PT full-length cDNAs defined in the specification, and for the detection
 XX PT and/or diagnosis of the abnormality of the proteins encoded by the
 XX PT full-length cDNAs -
 XX PS Claim 8; SEQ ID 18340; 2537pp + CD ROM; English.
 XX CC The present invention describes primer sets for synthesizing 5602
 XX CC full-length cDNAs defined in the specification. Where a primer set
 XX CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 XX CC to the complementary strand of a polynucleotide which comprises one of
 XX CC the 5602 nucleotide sequences defined in the specification, where the
 XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 XX CC of an oligonucleotide comprising a sequence complementary to the
 XX CC complementary strand of a polynucleotide which comprises a 5'-end
 XX CC sequence and an oligonucleotide comprising a sequence complementary to a
 XX CC polynucleotide which comprises a 3'-end sequence, where the
 XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
 XX CC the 3'-end sequence/3'-end sequence is selected from those defined in
 XX CC the specification. The primer sets can be used in antisense therapy and
 XX CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 XX CC particularly full-length cDNAs. The primers are also useful for the
 XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
 XX CC the full-length cDNAs. The primers allow obtaining of the full-length
 XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 XX CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 XX CC represent oligonucleotides, all of which are used in the exemplification
 XX CC of the present invention.
 XX SQ Sequence 394 AA;
 Query Match 50.5%; Score 48; DB 22; Length 394;
 Best Local Similarity 50.0%; Pred. No. 9.6;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 DDADQRIKQLEYFGNI 18
 I :|||:| :|||
 Db 265 DQSDQRIKLNHVGN 282
 RESULT 9
 AAG91310
 XX ID AAG91310 standard; Protein; 397 AA.
 XX AC AAG91310;
 XX DT 26-SEP-2001 (first entry)
 XX DE C glutamicum protein fragment SEQ ID NO: 5064.
 XX KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
 XX KW organic acid synthesis.
 XX OS Corynebacterium glutamicum.
 XX PR EP1108790-A2.
 XX PN

XX PD 20-JUN-2001.
 XX PF 18-DEC-2000; 2000EP-0127688.
 XX PR 16-DEC-1999; 99JP-0377484.
 XX PR 07-APR-2000; 2000JP-0159162.
 XX PR 03-AUG-2000; 2000JP-0280988.
 XX PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX DR WPI: 2001-376931/40.
 XX DR N-PSDB; AAH66529.
 XX PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 XX PT mutation point of a gene, measuring expression of a gene, analysing
 XX PT expression profile or pattern of a gene and identifying homologous gene
 XX PT -
 XX PS Claim 17; SEQ ID NO: 5064; 246pp + Sequence Listing; English.
 XX CC The present invention provides a number of nucleotide and protein
 XX CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 XX CC are useful for identifying the mutation point of a gene derived from a
 XX CC mutant of coryneform bacterium, measuring expression amount and
 XX CC analysing the expression profile or expression pattern of a gene derived
 XX CC from Coryneform bacterium, and identifying a homologue of a gene derived
 XX CC from coryneform bacterium. Coryneform bacteria are useful for producing
 XX CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
 XX CC particularly L-lysine. The present sequence is a protein described
 XX CC in the exemplification of the invention.
 XX CC Note: The sequence data for this patent did not form part of the printed
 XX CC specification, but was obtained in electronic format directly from the
 XX CC European Patent Office.
 XX SQ Sequence 397 AA;
 Query Match 50.5%; Score 48; DB 22; Length 397;
 Best Local Similarity 50.0%; Pred. No. 9.7;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 2 DADQRIKQLEYFGN 17
 ||| :| :|||
 Db 138 DAQOSLLDRMEAYFGN 153
 RESULT 10
 AAG01240
 XX ID AAG01240 standard; Protein; 650 AA.
 XX AC AAG01240;
 XX DT 13-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #1231.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PR 30-MAR-2001; 2001WO-US08631.
 XX PR 31-MAR-2000; 2000US-0540217.
 XX PR 23-AUG-2000; 2000US-0649167.
 XX PN


```

XX PR 16-SEP-1999; 99US-0397550.
XX PA (WARN ) WARNER LAMBERT CO.
XX PI Brown JP, Bertelli F;
XX DR WPI: 2001-235262/24.
XX DR N-PSDB; AAS01404.
XX PT Calcium channel alpha2delta subunits, useful in e.g. SPA assays,
XX PT Flashplate assays, Nickel Flashplate assays, Filter binding assays or
XX PT Wheat Germ Lectin Flashplate assays -
XX PS Claim 22; Page 58-61; 160pp; English.
XX CC The present sequence represents human secreted calcium channel
XX CC alpha2delta subunit #1 which is soluble and retains the functional
XX CC characteristics of the full length or wild type alpha2delta subunit
XX CC (AAU01025) from which it is derived. The invention relates to truncated
XX CC alpha2delta-2, alpha2delta-3 or alpha2delta-4 subunit soluble proteins
XX CC which retain their affinity for radioactively labelled gabapentin. The
XX CC alpha2delta subunit is 1 of the components of the heteromultimeric
XX CC voltage-dependent calcium channel (VDCC) complexes present in neuronal
XX CC and non-neuronal tissues including heart and skeletal muscle. Numerous
XX CC soluble forms of the human calcium channel alpha2delta subunits
XX CC (AAU01014-AAU01024 and AAU01032-AAU01038) and 5 soluble forms of the
XX CC porcine calcium channel alpha2delta subunits (AAU01027-AAU01031) are
XX CC described. The secreted soluble alpha2delta subunit may be used in assays
XX CC e.g. scintillation proximity assay (SPA), flashplate, nickel flashplate,
XX CC filter binding or wheat germ lectin flashplate assays to detect or
XX CC measure the binding or interaction of a ligand (e.g. gabapentin,
XX CC L-Norleucine, L-Allo-Isoleucine, L-methionine, L-Leucine, L-Isoleucine,
XX CC L-Valine, Spermine and/or L-Phenylalanine) of a calcium channel
XX CC alpha2delta subunit.
XX SQ Sequence 1062 AA;
    Query Match 47.4%; Score 45; DB 22; Length 1062;
    Best Local Similarity 50.0%; Pred. No. 99;
    Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 7 IIRKLEYFNGNI 18
    :||:||||:
Db 1007 VMKOTQYFGSV 1018

RESULT 15
AAB62238
XX ID AAB62238.standard; Protein; 1062 AA.
XX AC AAB62238;
XX DT 11-JUN-2001 (first entry)
XX DE Human calcium channel subunit alpha2delta-1.
XX KW Calcium channel alpha2delta subunit; alpha2delta-1; cerebral cortex;
XX KW nervous system disorder; pain; epilepsy; anxiety; human.
XX OS Homo sapiens.
XX PN WO200120336-A2.
XX PD 22-MAR-2001.
XX PF 18-SEP-2000; 2000WO-EP09136.
XX PR 16-SEP-1999; 99US-0397549.
XX PA (WARN ) WARNER LAMBERT CO.
XX PI Bertelli F, Brown JP, Dissanayake V, Suman-Chauhan N, Gee NS;

```

```

XX DR WPI: 2001-257902/26.
XX DR N-PSDB; AAF57541.
XX PT Competitive binding assay for screening ligands which bind a cerebral
XX PT cortical voltage-dependent calcium channel alpha2delta-1 subunit,
XX PT where the ligands identified are useful for treating disorders of the
XX PT nervous system, including pain -
XX PS Disclosure; Page 59-62; 158pp; English.
XX CC The invention relates to a new method for screening ligands which bind a
XX CC cerebral cortical voltage-dependent calcium channel alpha2delta subunit,
XX CC preferably alpha2delta-1 subunit. The method comprises contacting a
XX CC secreted soluble recombinant alpha2delta-1 subunit with a ligand of
XX CC interest and a labelled compound which binds the subunit, followed by
XX CC measuring the level of binding of the labelled compound to alpha2delta-1
XX CC subunit. The method is useful for screening ligands, preferably
XX CC biologically active products that modulate a nervous system function,
XX CC which bind a cerebral cortical voltage-dependent calcium channel
XX CC alpha2delta-1 subunit. The ligands identified by the method are useful
XX CC for treating disorders of the nervous system, including pain, epilepsy
XX CC and anxiety. The present sequence represents a human calcium channel
XX CC subunit alpha2delta-1.
XX SQ Sequence 1062 AA;
    Query Match 47.4%; Score 45; DB 22; Length 1062;
    Best Local Similarity 50.0%; Pred. No. 99;
    Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
QY 7 IIRKLEYFNGNI 18
    :||:||||:
Db 1007 VMKOTQYFGSV 1018

Search completed: April 23, 2003, 13:27:16
Job time : 29.5169 secs

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:34:48 ; Search time 20.0225 Seconds
(without alignments)
185.234 Million cell updates/sec

Title: US-09-836-073-17

Perfect score: 95
Sequence: 1 DDADQRIKQLEYFGNI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 9297

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	34.7	25	8 Q9XJB7	Q9XJB7 aureobrya
2	33	34.7	25	8 Q8W680	Q8W680 bacterioph
3	32	33.7	11	10 O82070	O82070 triticum ae
4	29	30.5	16	2 Q9R963	Q9R963 helicobacte
5	29	30.5	23	13 Q9PRR5	Q9PRR5 anser anser
6	28	29.5	22	4 Q96P08	Q96P08 homo sapien
7	27	28.4	17	7 Q19716	Q19716 homo sapien
8	27	28.4	24	5 Q95SB6	Q95SB6 drosophila
9	26	27.4	15	9 Q9MBH6	Q9MBH6 streptococc
10	26	27.4	16	2 P82955	P82955 acinetobact
11	26	27.4	23	2 Q9F9U1	Q9F9U1 pseudomonas
12	26	27.4	25	2 Q9R527	Q9R527 escherichia
13	26	27.4	25	8 Q9TGB8	Q9TGB8 alnus crisp
14	26	27.4	25	8 Q9TGB7	Q9TGB7 alnus gluti
15	26	27.4	25	8 Q9TGB6	Q9TGB6 alnus marit
16	26	27.4	25	8 Q9TGB5	Q9TGB5 betula alle

17	26	27.4	25	8 Q9TGB4	Q9TGB4 betula glan
18	26	27.4	25	8 Q9TGB3	Q9TGB3 betula papy
19	26	27.4	25	8 Q9TGB2	Q9TGB2 betula verr
20	26	27.4	25	8 Q9TGB1	Q9TGB1 betula pube
21	26	27.4	25	8 Q9TGB0	Q9TGB0 corylus ave
22	26	27.4	25	8 Q9TGA9	Q9TGA9 corylus col
23	26	27.4	25	8 Q9TGA8	Q9TGA8 corylus cor
24	26	27.4	25	8 Q9TGA7	Q9TGA7 ostrya virg
25	26	27.4	25	8 Q9TGA6	Q9TGA6 quercus rub
26	26	27.4	25	8 Q9TGB3	Q9TGB3 carpinus ca
27	26	27.4	25	15 Q9WRF2	Q9WRF2 human immun
28	26	27.4	25	15 Q9TQO1	Q9TQO1 human immun
29	26	27.4	25	15 Q9TQO2	Q9TQO2 human immun
30	26	27.4	25	15 Q9TQO4	Q9TQO4 human immun
31	25	26.3	18	2 Q57323	Q57323 yersinia ps
32	25	26.3	20	6 Q9TR34	Q9TR34 ovis aries
33	25	26.3	20	6 Q9TR33	Q9TR33 canis famil
34	25	26.3	20	9 Q37868	Q37868 bacterioph
35	25	26.3	24	2 Q05616	Q05616 staphylococ
36	25	26.3	24	15 Q99DC3	Q99DC3 human immun
37	25	26.3	24	15 Q99DA3	Q99DA3 human immun
38	25	26.3	25	15 Q72001	Q72001 human immun
39	25	26.3	25	15 Q9TQ99	Q9TQ99 human immun
40	25	26.3	25	15 Q8QDY1	Q8QDY1 human immun
41	25	26.3	25	15 Q8QDX7	Q8QDX7 human immun
42	24	25.3	14	4 Q16118	Q16118 homo sapien
43	24	25.3	18	4 Q16028	Q16028 homo sapien
44	24	25.3	18	11 Q64134	Q64134 rattus sp.
45	24	25.3	19	1 Q50844	Q50844 methanococc

ALIGNMENTS

RESULT 1

Q9XJB7	Q9XJB7	PRELIMINARY;	PRT;	25 AA.
AC	Q9XJB7			
DT	01-NOV-1999 (T-REMBLrel. 12, Created)			
DT	01-NOV-1999 (T-REMBLrel. 12, Last sequence update)			
DT	01-DEC-2001 (T-REMBLrel. 19, Last annotation update)			
DE	Ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit (Fragment).			
DE	(Fragment).			
GN	RBCS.			
OS	Aureobrya lagunensis.			
OG	Chloroplast.			
OC	Eukaryota; stramenopiles; Pelagophyceae; Aureobrya.			
OX	NCBI_TaxID=44058;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CCMP1681, CCMP1509, AND CCMP1510;			
RA	Bailey J.C., Andersen R.A.;			
RT	"Analysis of clonal cultures of the brown tide algae Aureococcus and Aureobrya (Pelagophyceae) using 18S rRNA, rbcL and rubisco spacer sequences."			
RL	Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF117782; AAD39106.1;			
DR	EMBL; AF117780; AAD39104.1;			
DR	EMBL; AF117781; AAD39105.1;			
DR	InterPro; IPR000894; RuBisCO_small.			
DR	Pfam; PF00101; RuBisCO_small; 1.			
KW	Chloroplast.			
FT	NON_TPR			
SQ	SEQUENCE 25 AA; 2972 MW; 61A96BD83DB94118 CRC64;			

Query Match 34.7%; Score 33; DB 8; Length 25;
Best Local Similarity 53.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DDADQRIKQLEY 13

Db 13 DLTDQIRKQVEY 25

```

RESULT 2
Q8W680
ID Q8W680 PRELIMINARY; PRT; 25 AA.
AC Q8W680;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE DNA helix destabilizing protein (Fragment).
GN G32.
OS Bacteriophage RB49.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC 4-like viruses.
OX NCBI_TaxID=50948;
RN [1]
RP SEQUENCE FROM N.A.
RA Desplats C., Dez C., Tetart F., Eleaume H., Krusch H.M.;
RT "A snapshot of the genome of the pseudo T-even bacteriophage.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY051176; AAL15126.1; -
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 3060 MW; 6CA91FCC0C4D7D02 CRC64;

Query Match 34.7%; Score 33; DB 9; Length 25;
Best Local Similarity 33.3%; Pred. No. 2.1e+02;
Matches 6; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 DDADQRIIKQLEYIFGNI 18
| : : : : |
DB 4 DKTEWOLIKRASYANI 21

RESULT 3
O82070
ID O82070 PRELIMINARY; PRT; 11 AA.
AC O82070;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE RNA polymerase (EC 2.7.7.6) (Fragment).
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. CHINESE SPRING;
RA Young D.A., Allen R., Harvey A.J., Lonsdale D.M.;
RT "Characterization of a gene encoding a single-subunit RNA polymerase
RT from maize which is alternatively spliced.";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ005344; CAA06489.1; -
KW Nucleotidyltransferase; Transferase.
FT NON_TER 1
FT NON_TER 11
SQ SEQUENCE 11 AA; 1329 MW; CD96344923240AB2 CRC64;

Query Match 33.7%; Score 32; DB 10; Length 11;
Best Local Similarity 40.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 DDADQRIIKQ 10
| : : : : |
DB 2 DEVDRKLVKQ 11

RESULT 4
Q9R963
ID Q9R963 PRELIMINARY; PRT; 16 AA.
AC Q9R963;
DT 01-MAY-2000 (TRENBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE FEPC (Fragment).
GN FEPC.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F31;
RX MEDLINE=98453456; PubMed=9780360;
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,
RA Kuriyama M.;
RT "Full-length sequence analysis of the vacA gene from cytotoxic and
RT noncytotoxic Helicobacter pylori.";
RL J. Infect. Dis. 178:1391-1398(1998).
DR EMBL; AF049623; AAD04263.1; -
FT NON_TER 1
FT NON_TER 16
SQ SEQUENCE 16 AA; 1938 MW; 3C33D403AEE61428 CRC64;

Query Match 30.5%; Score 29; DB 2; Length 16;
Best Local Similarity 60.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 IKQLEYIFGN 17
| : : : : |
DB 5 IPQVEYIAPN 14

RESULT 5
Q9PRR5
ID Q9PRR5 PRELIMINARY; PRT; 25 AA.
AC Q9PRR5;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE Apolipoprotein CA (Fragment).
OS Anser anser (domestic goose).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anser.
OX NCBI_TaxID=8843;
RN [1]
RP SEQUENCE
RX MEDLINE=96128192; PubMed=8536707;
RA Hermier D., Sellier N., Rousselot-Pailley D., Forgez P.;
RT "Characterization of apolipoproteins B-100, AI and C from plasma
RT lipoprotein in the goose, Anser anser. Evidence for a genetic
RT polymorphism in Apoc-like apolipoproteins.";
RL Eur. J. Biochem. 234:586-591(1995).
SQ SEQUENCE 25 AA; 2804 MW; AC4EA982117B346E CRC64;

Query Match 30.5%; Score 29; DB 13; Length 25;
Best Local Similarity 46.2%; Pred. No. 9.3e+02;
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 DDADQRIIKQLEY 13
| : : : : |
DB 4 DDTDTVWVKVQY 16

RESULT 6
Q96P08
ID Q96P08 PRELIMINARY; PRT; 22 AA.
AC Q96P08;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Upstream ORF.
GN BRE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21534249; PubMed=11676476;
RA Ching A.K.K., Li P.S., Li Q., Chan B.C.L., Chan J.Y.H., Lim P.L.,
RA Pang J.C.S., Chui Y.L.;
RT "Expression of human BRE in multiple isoforms";
RL Biochem. Biophys. Res. Commun. 288:535-545(2001).
DR EMBL; AF420603; AAL17815.1; -;
SQ SEQUENCE 22 AA; 2402 MW; E2BDD3D056462A8 CRC64;

Query Match 29.5%; Score 28; DB 4; Length 22;
Best Local Similarity 60.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 KQLEYFGNI 18
| | | | |
Db 6 KNTTEQFGNI 15

RESULT 7

OL19716 PRELIMINARY; PRT; 17 AA.

AC OL19716;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Human histocompatibility system hla-dr heavy chain (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=82197531; PubMed=6952207;
RA Lee J.S., Trowsdale J., Bodmer W.F.;
RT "cdna clones coding for the heavy chain of human hla-dr antigen";
RL Proc. Natl. Acad. Sci. U.S.A. 79:545-549(1982).
DR EMBL; J00193; AAA36272.1; -;
FT NON_TER 1 1
FT NON_TER 17 17
SQ SEQUENCE 17 AA; 2058 MW; 4874E328EE648F54 CRC64;

Query Match 28.4%; Score 27; DB 7; Length 17;
Best Local Similarity 36.4%; Pred. No. 1.3e+03;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 DQRIKQLEY 14
: : | | | |
Db 3 EEHVIIQAEFY 13

RESULT 8

Q95SB6 PRELIMINARY; PRT; 24 AA.

AC Q95SB6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE GM03811P.
GN CG11779.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Prorygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.

RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan B., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,

RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY060874; AAL28422.1; -;
DR Flybase; FBgn0038683; CG11779.
SQ SEQUENCE 24 AA; 3105 MW; AAFB07E185CA5104 CRC64;

Query Match 28.4%; Score 27; DB 5; Length 24;
Best Local Similarity 50.0%; Pred. No. 1.9e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 RIHKQLEYF 15
| | | | |
Db 3 RKIKILQFY 12

RESULT 9

Q9MBH6 PRELIMINARY; PRT; 15 AA.

AC Q9MBH6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Integrase (Fragment).
GN INT.
OS Streptococcus pneumoniae bacteriophage MM1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=120574;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20392133; PubMed=10933687;
RA Gindreau E., Lopez R., Garcia P.;
RT "MM1, a temperate bacteriophage of the type 23F Spanish/USA
multiresistant epidemic clone of Streptococcus pneumoniae: structural
analysis of the site-specific integration system";
RL J. Virol. 74:7803-7813(2000).
DR EMBL; AJ400630; CAB96618.1; -;
FT NON_TER 1 1
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1810 MW; D9E4B2ABE94E3543 CRC64;

Query Match 27.4%; Score 26; DB 9; Length 15;
Best Local Similarity 20.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 1 DDADQRIKQLEYF 15
: | : : : : | |
Db 1 EDMEDKLVNKLDITF 15

RESULT 10

P82955 PRELIMINARY; PRT; 16 AA.

AC P82955;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Alkyl hydroperoxide reductase subunit F (EC 1.6.4.-)
DE (Fragment).
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=471;
RN [1]
RP SEQUENCE, AND INDUCTION.
RC STRAIN=69-V;
RX MEDLINE=21318969; PubMed=11425483;
RA Benndorf D., Loffhagen N., Babel W.;
RT "Protein synthesis patterns in Acinetobacter calcoaceticus induced by
phenol and catechol show specificities of responses to chemostress";
RL FEMS Microbiol. Lett. 200:247-252(2001).
CC -1- FUNCTION: SERVES TO PROTECT THE CELL AGAINST DNA DAMAGE BY ALKYL
HYDROPEROXIDES. IT CAN USE EITHER NADH OR NADPH AS ELECTRON DONOR
FOR DIRECT REDUCTION OF REDOX DYES OR OF ALKYL HYDROPEROXIDES WHEN
COMBINED WITH THE AHPC PROTEIN (BY SIMILARITY).

CC -1- COFACTOR: FAD (BY SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- INDUCTION: BY OXIDATIVE STRESS AND CATECHOL. INDUCED VERY WEAKLY
 CC BY HEAT SHOCK.
 CC -1- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
 CC OXIDOREDUCTASES CLASS-II.
 DR InterPro: IPR00103; Pyridine_redox_2.
 DR PROSITE: PS00573; PYRIDINE_REDOX_2; PARTIAL.
 KW Redox-active center; Oxidoreductase; NADP; NAD; Flavoprotein; FAD;
 KW Heat shock.
 FT NON_TER 16 16
 SQ SEQUENCE 16 AA; 1928 MW; E070754AFCBED15 CRC64;

Query Match 27.4%; Score 26; DB 2; Length 16;
 Best Local Similarity 54.5%; Pred. No. 1.8e+03;
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 DQRIKQLEYFGN 14
 || || || ||
 DB 3 DQNIXTQLXAY 13

RESULT 11
 Q9F9U1 PRELIMINARY; PRT; 23 AA.
 AC Q9F9U1
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
 DE Coenzyme PQQ synthesis protein A.
 GN PQQA.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BC-2;
 RA Chang C.-H., Herrick J.B., Okinaka R.T., Brainard J.B.,
 RA Terwilliger T.C.;
 RT "Identification and characterization of genes activated by 2-
 RT chloroethanol in Pseudomonas stutzeri BC-2.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF176640; AAG09250.1;
 SQ SEQUENCE 23 AA; 2809 MW; 80815999B7EBCC47 CRC64;

Query Match 27.4%; Score 26; DB 2; Length 23;
 Best Local Similarity 42.9%; Pred. No. 2.6e+03;
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 DQRIKQLEYFGN 17
 || || || ||
 DB 9 DLRIQFVTLXFAN 22

RESULT 12
 Q9R5Z7 PRELIMINARY; PRT; 25 AA.
 AC Q9R5Z7
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
 DE Putative colonization factor (Fragment).
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94041642; PubMed=7901165;
 RA Viboud G.I., Binsztain N., Svennerholm A.M.;
 RT "A new fimbrial putative colonization factor, PCF020, in human
 RT enterotoxigenic Escherichia coli.";

RL Infect. Immun. 61:5190-5197(1993).
 SQ SEQUENCE 25 AA; 2621 MW; 7959C9E0EA955C66 CRC64;

Query Match 27.4%; Score 26; DB 2; Length 25;
 Best Local Similarity 26.7%; Pred. No. 2.9e+03;
 Matches 4; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 4 DQRIKQLEYFGN 18
 | | | | | | | |
 DB 6 DDSXXATLDFTGNV 20

RESULT 13

Q9TGB8 PRELIMINARY; PRT; 25 AA.
 AC Q9TGB8;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Ribosomal protein small 3 (Fragment).
 GN RPS3.
 OS Alnus crispa.
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fagales; Betulaceae; Alnus.
 OX NCBI_TaxID=3518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99261656; PubMed=10331271;
 RA Laroche J., Bousquet J.;
 RT "Evolution of the mitochondrial rps3 intron in perennial and annual
 RT angiosperms and homology to nad5 intron 1.";
 RL Mol. Biol. Evol. 16:441-452(1999).
 DR EMBL: AF080075; AAD50062.1;
 DR InterPro: IPR001351; Ribosomal_S3.
 DR Pfam: PF00417; Ribosomal_S3_N; 1.
 KW Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 27.4%; Score 26; DB 8; Length 25;
 Best Local Similarity 62.5%; Pred. No. 2.9e+03;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 LEYFGNI 18
 | | | | |
 DB 17 LRSYFGSI 24

RESULT 14

Q9TGB7 PRELIMINARY; PRT; 25 AA.
 AC Q9TGB7;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Ribosomal protein small 3 (Fragment).
 GN RPS3.
 OS Alnus glutinosa (Alder).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fagales; Betulaceae; Alnus.
 OX NCBI_TaxID=3517;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99261656; PubMed=10331271;
 RA Laroche J., Bousquet J.;
 RT "Evolution of the mitochondrial rps3 intron in perennial and annual
 RT angiosperms and homology to nad5 intron 1.";
 RL Mol. Biol. Evol. 16:441-452(1999).

DR EMBL; AF080076; AAD50063.1; -
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 27.4%; Score 26; DB 8; Length 25;
Best Local Similarity 62.5%; Pred. No. 2.9e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 LEYFEGNI 18
| | | | |
Db 17 LRSYFGSI 24

RESULT 15
Q9TGB6 PRELIMINARY; PRT; 25 AA.
AC Q9TGB6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Ribosomal protein small 3 (Fragment).
GN RPS3.
OS Alnus maritima.
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Alnus.
OX NCBI_TaxID=21015;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99261656; PubMed=10331271;
RA Laroche J., Bousquet J.;
RT "Evolution of the mitochondrial rps3 intron in perennial and annual
angiosperms and homology to nad5 intron 1.";
RL Mol. Biol. Evol. 16:441-452(1999).
DR EMBL; AF080077; AAD50064.1; -
DR InterPro; IPR001351; Ribosomal_S3.
DR Pfam; PF00417; Ribosomal_S3_N; 1.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 27.4%; Score 26; DB 8; Length 25;
Best Local Similarity 62.5%; Pred. No. 2.9e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 LEYFEGNI 18
| | | | |
Db 17 LRSYFGSI 24

Search completed: April 23, 2003, 13:47:18
Job time : 22.0225 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:33:13 ; Search time 4.65169 Seconds
(without alignments)
160.495 Million cell updates/sec

Title: US-09-836-073-17

Perfect score: 95
Sequence: 1 DDADQRIKQLEYFGNI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1520

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	28.4	24	1 DHE3_PYRWO	Q09115 pyrococcus
2	24	25.3	20	1 THIO_CANFA	P99505 canis famill
3	24	25.3	20	1 YPRB_SERMA	P22581 serratia ma
4	23	24.2	10	1 MALE_KLEPN	Q05564 klebsiella
5	23	24.2	19	1 FIBB_VULVU	P14482 vulpes vulp
6	22	23.2	14	1 TKNM_RANMA	P40951 rana margar
7	22	23.2	16	1 AU21_LITRA	P82388 litoria ran
8	22	23.2	16	1 AU25_LITRA	P82392 litoria ran
9	22	23.2	16	1 BAIT_EUBSP	P32371 eubacterium
10	22	23.2	19	1 HBB2_UROHA	P18992 uromastix h
11	22	23.2	22	1 CYSP_TRIVA	P33404 trichomonas
12	22	23.2	23	1 NUO5_SOLTU	P80262 solanum tub
13	22	23.2	24	1 DLAA_AMEYE	P80413 anycolatops
14	22	23.2	24	1 LCA_FELCA	P37154 felis silve
15	22	23.2	24	1 RS13_THETH	P80377 thermus the
16	22	23.2	25	1 TBA_IETPE	P10489 leptomonas
17	21	22.1	13	1 FIBB_RABIT	P14478 oryctolagus
18	21	22.1	18	1 FIBB_ANAPL	P12802 anas platyr
19	21	22.1	18	1 PHTP_PSESE	P25271 pseudalecia
20	21	22.1	23	1 AU43_LITAU	P82399 litoria aur
21	21	22.1	24	1 LPA2_BACSU	P08497 bacillus su
22	21	22.1	24	1 PQQA_PSEFL	P55171 pseudomonas
23	21	22.1	25	1 ACP_ACICA	P80916 acinetobact
24	21	22.1	25	1 DNAB_MYCCA	P71500 mycoplasma
25	21	22.1	25	1 PRLA_ACHLY	P27459 achromobact
26	21	22.1	25	1 V23K_WSSW	Q09124 lactococcus
27	20	21.1	12	1 SCRK_LACLA	P82005 white spot
28	20	21.1	14	1 RS19_PRUAP	Q44160 prunus arme
29	20	21.1	16	1 MLB_SOUAC	P01207 squalus aca
30	20	21.1	16	1 UVSX_BPT6	Q06728 bacterioph
31	20	21.1	17	1 UF36_UPEMJ	P82043 escherichia m
32	20	21.1	18	1 FMF1_ECOLI	P20860 escherichia
33	20	21.1	19	1 CH10_CLOPA	P81338 clostridium

34 20 21.1 19 1 NUO6_SOLTU P80729 solanum tub
35 20 21.1 22 1 UVSX_BPT2 Q06727 bacterioph
36 20 21.1 23 1 AU42_LITRA P82398 litoria ran
37 20 21.1 23 1 PRO3_DACGL P18690 dactylis gl
38 20 21.1 23 1 SODM_RANCA P36215 rana catesb
39 20 21.1 24 1 ACHA_ELEEL P09688 electrophor
40 20 21.1 24 1 DNAB_STRAG P55694 streptococc
41 20 21.1 24 1 LEC_CROST P16351 crotalaria
42 20 21.1 24 1 PL13_PLETR P36987 plectreury
43 20 21.1 24 1 PQQA_ACICA P27532 acinetobact
44 20 21.1 25 1 NEUU_PIG P34964 sus scrofa
45 19 20.0 8 1 AL17_CARMA P81820 carcinus ma

ALIGNMENTS

RESULT 1

DHE3_PYRWO STANDARD; PRT; 24 AA.
ID Q09115;
AC Q09115;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Glutamate dehydrogenase (EC 1.4.1.3) (GDH) (Fragment).
GN GDHA OR GDH.
OS Pyrococcus woesei.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2262;
RN [1]
RP SEQUENCE.
RX MEDLINE=94010338; PubMed=8406037;
RA Eagen R.I.L., Geerling A.C.M., Waldkoetter K., Antranikian G.,
de Vos W.M.;
RT "The glutamate dehydrogenase-encoding gene of the hyperthermophilic
archaeon Pyrococcus furiosus: sequence, transcription and analysis of
the deduced amino acid sequence."
RL Gene 132:143-148(1993).
CC -|- CATALYTIC ACTIVITY: L-glutamate + H(2)O + NAD(P)(+) -> 2-
oxoglutarate + NH(3) + NAD(P)H.
CC -|- SUBUNIT: Homohexamer (By similarity).
CC -|- SIMILARITY: BELONGS TO THE GLU/LEU/PHE/VAL DEHYDROGENASES FAMILY.
DR HSP; P80319; 1GTM.
DR InterPro; IPR001625; GLFV_Dh.
DR PROSITE; PS00074; GLFV_DEHYDROGENASE; PARTIAL.
KW Oxidoreductase; NAD; NADP.
FT NON_TER 24
SQ SEQUENCE 24 AA; 2882 MW; 67FD81A4C45DBBC7 CRC64;

Query Match 28.4%; Score 27; DB 1; Length 24;
Best Local Similarity 54.5%; Pred. No. 2.3e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 DDADQRIKQLE 12
| : |||||
Db 4 DPYEIVIKQLE 14

RESULT 2

THIO_CANFA STANDARD; PRT; 20 AA.
ID THIO_CANFA
AC P99505;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Thioredoxin (Fragment).
GN TXN.

OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]

```

RC SEQUENCE.
RA TISSUE=Heart;
RL Dunn M.J., Wheeler C.H.;
RL Submitted (AUG-1997) to the SWISS-PROT data bank.
CC -!- FUNCTION: Participates in various redox reactions through the
CC reversible oxidation of its active center dithiol to a disulfide
CC and catalyzes dithiol-disulfide exchange reactions.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY.
DR HSP; P10599; IERW.
DR HSC-2DPAGE; P99505; DOG.
DR InterPro: IPR000063; Thiorod.
DR PROSITE: PS00194; THIOREDOXIN; PARTIAL.
KW Redox-active center; Electron transport.
FT UNSURE 6
FT NON_TER 20
SQ SEQUENCE 20 AA; 2287 MW; A06991862EB1B6A6 CRC64;

Query Match 25.3%; Score 24; DB 1; Length 20;
Best Local Similarity 37.5%; Pred. No. 6e+02;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 8 IKQLEYF 15
DB 1 VQKIEFKY 8

RESULT 3
YPRB_SERMA STANDARD; PRT; 20 AA.
AC P22581;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in PROB 5/region (Fragment).
OS Serratia marcescens.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sr41;
RA Omori K., Suzuki S., Imai Y., Komatsubara S.;
RT "Analysis of the Serratia marcescens proBA operon and feedback
RT control of proline biosynthesis.";
RL J. Gen. Microbiol. 137:509-517(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90351; BAA14363.1; -.
DR EMBL; X53086; CAA37253.1; -.
DR PIR; S11643; S11643.
DR PIR; C49753; C49753.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 20 AA; 2248 MW; 4DD7777735276674 CRC64;

Query Match 25.3%; Score 24; DB 1; Length 20;
Best Local Similarity 44.4%; Pred. No. 6e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 DDADQRIIK 9
DB 9 DDFKEKLIK 17

```

```

RESULT 4
MALE_KLEPN STANDARD; PRT; 10 AA.
ID MALE_KLEPN
AC Q05564;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Maltose-binding periplasmic protein (Maltodextrin-binding protein)
DE (MBP) (Fragment).
GN MALE.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1033-5p14 / KAY2026;
RX MEDLINE=93211295; PubMed=8459773;
RA Bachelier S., Perrin D., Hofnung M., Gilson E.;
RT "Bacterial interspersed mosaic elements (BIMEs) are present in the
RT genome of Klebsiella.";
RL Mol. Microbiol. 7:537-544(1993).
CC -!- FUNCTION: MALE IS INVOLVED IN THE HIGH-AFFINITY MALTOSE MEMBRANE
CC TRANSPORT SYSTEM. INITIAL RECEPTOR FOR THE ACTIVE TRANSPORT OF AND
CC CHEMOTAXIS TOWARD MALTOOLIGOSACCHARIDES.
CC -!- SUBCELLULAR LOCATION: Periplasmic.
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
CC PROTEIN FAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X68329; CAA48406.1; -.
DR InterPro: IPR000567; SBP_bac1.
DR PROSITE: PS01037; SBP_BACTERIAL_1; PARTIAL.
KW Transport; Sugar transport; Periplasmic.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1159 MW; 8FD8DC4415A6DDDA CRC64;

Query Match 24.2%; Score 23; DB 1; Length 10;
Best Local Similarity 62.5%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 DADQRIIK 9
DB 3 DAQSRITK 10

RESULT 5
FIBB_VULVU STANDARD; PRT; 19 AA.
ID FIBB_VULVU
AC P14482;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
GN FGB.
OS Vulpes vulpes (Red fox).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Vulpes.
OX NCBI_TaxID=9627;
RN [1]
RP SEQUENCE.
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET

```

```

CC AGGREGATION.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR INTERPRO: IPR002181; Fibrinogen_C.
DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 19 FIBRINOPEPTIDE B.
FT NON_TER 19
FT SEQUENCE 19 AA; 2305 MW; 0B0D756C24A1F2C0 CRC64;

Query Match 24.2%; Score 23; DB 1; Length 19;
Best Local Similarity 27.3%; Pred. No. 8.3e+02;
Matches 3; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 2 DADQRIKOLE 12
DB 7 DEERIVSTVD 17

RESULT 6
TKNM_RANMA STANDARD; PRT; 14 AA.
AC P40951;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ranamargarin.
OS Rana margaratae (Chinese frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX NCBI_TaxID=121156;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=90026852; PubMed=2803524;
RA Tang Y.Q., Tian S.H., Wu S.X., Hua J.C., Wu G.F., Zhao E.M., Lu Y.A.,
RA Zhu Y.Q., Zou G., Tsou K.;
RT "Isolation and structure of ranamargarin, a new tachykinin from the
RT skin of Chinese frog Rana margaratae."
RL Sci. China, B, Chem. Life Sci. Earth Sci. 32:570-579(1989).
RN [2]
RP SYNTHESIS.
RX MEDLINE=90253600; PubMed=2340087;
RA Lu Y.A., Peng J.L., Zhu Y.Q., Wu S.X., Tang Y.Q., Tian S.H., Zou G.;
RT "Synthesis and biological activity of a new frog skin peptide,
RT ranamargarin."
RL Sci. China, B, Chem. Life Sci. Earth Sci. 33:170-177(1990).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR InterPro: IPR002040; Tachykinin.
DR PROSITE: PS00267; TACHYKININ: 1.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 14 14 AMIDATION.
FT SEQUENCE 14 AA; 1617 MW; D4593AE408C3673D CRC64;

Query Match 23.2%; Score 22; DB 1; Length 14;
Best Local Similarity 55.6%; Pred. No. 8.6e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 DDADQRIK 9
DB 1 DDASDRKK 9

RESULT 7
AU21_LITRA

```

```

ID AU21_LITRA STANDARD; PRT; 16 AA.
AC P82388;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aurein 2.1 [Contains: Aurein 2.1.1].
OS Litoria raniformis (Southern bell frog), and
OS Litoria aurea (Green and golden bell frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=116057, 8371;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Granular dorsal gland;
RX MEDLINE=20408845; PubMed=10951191;
RA Rozek T., Wegener K.L., Bowie J.H., Oliver I.N., Carver J.A.,
RA Wallace J.C., Tyler M.J.;
RT "The antibiotic and anticancer active aurein peptides from the
RT Australian bell frogs Litoria aurea and Litoria raniformis the
RT solution structure of aurein 1.2."
RL Eur. J. Biochem. 267:5330-5341(2000).
CC -1- FUNCTION: AUREIN 2.1 HAS ANTIMICROBIAL ACTIVITY AGAINST B.CEREUS,
CC L.LACTIS, L.INNOCUA, M.LUTEUS, S.EPIDERMIDIS AND S.UBERISAND.
CC PROBABLY ACTS BY DISTURBING MEMBRANE FUNCTIONS WITH ITS
CC AMPHIPATHIC STRUCTURE.
CC -1- SUBCELLULAR LOCATION: SECRETED.
KW Amidation; Antibiotic.
FT PEPTIDE 1 16 AUREIN 2.1.
FT PEPTIDE 3 16 AUREIN 2.1.1.
FT MOD_RES 16 16 AMIDATION.
FT SEQUENCE 16 AA; 1616 MW; 1D9A5DAD84D240F9 CRC64;

Query Match 23.2%; Score 22; DB 1; Length 16;
Best Local Similarity 33.3%; Pred. No. 1e+03;
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 7 IKOLEYFYFGNI 18
DB 5 IVKVVGFAGFSL 16

RESULT 8
AU25_LITRA STANDARD; PRT; 16 AA.
AC P82392;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aurein 2.5.
OS Litoria raniformis (Southern bell frog), and
OS Litoria aurea (Green and golden bell frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=116057, 8371;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Granular dorsal gland;
RX MEDLINE=20408845; PubMed=10951191;
RA Rozek T., Wegener K.L., Bowie J.H., Oliver I.N., Carver J.A.,
RA Wallace J.C., Tyler M.J.;
RT "The antibiotic and anticancer active aurein peptides from the
RT Australian bell frogs Litoria aurea and Litoria raniformis the
RT solution structure of aurein 1.2."
RL Eur. J. Biochem. 267:5330-5341(2000).
CC -1- FUNCTION: ANTIMICROBIAL ACTIVITY AGAINST B.CEREUS, L.LACTIS,
CC L.INNOCUA, M.LUTEUS, S.AUREUS, AND S.EPIDERMIDIS. PROBABLY ACTS BY
CC DISTURBING MEMBRANE FUNCTIONS WITH ITS AMPHIPATHIC STRUCTURE.
CC SHOWS ANTICANCER ACTIVITY.
CC -1- SUBCELLULAR LOCATION: SECRETED.
KW Amidation; Antibiotic.
FT MOD_RES 16 16 AMIDATION.

```

SQ SEQUENCE 16 AA; 1650 MW; 1D9A5DADB4DAE2F9 CRC64;

Query Match 23.2%; Score 22; DB 1; Length 16;
Best Local Similarity 33.3%; Pred. No. 1e+03; 3; Indels 0; Gaps 0;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 7 IIKQLEYFNGI 18
I::: I:::
Db 5 IVKVVGFAGSL 16

RESULT 9
BAIL_EUBSP STANDARD; PRT; 16 AA.
AC P32371;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bile acid-inducible operon protein I (Fragment).
GN BAIL.
OS Eubacterium sp. (strain VPI 12708).
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=29347;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9325945; PubMed=8491719;
RA Franklund C.V., Baron S.F., Hylemon P.B.;
RT "Characterization of the baiH gene encoding a bile acid-inducible
NADH:flavin oxidoreductase from Eubacterium sp. strain VPI 12708.";
RL J. Bacteriol. 175:3002-3012(1993).
CC -|- PATHWAY: Bile acid catabolism.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U57489; AAC45418.1; -
DR PIR; D36912; D36912.
KW Bile acid catabolism.
FT NON_TER 16
SQ SEQUENCE 16 AA; 1721 MW; B9BAAE7B8746E4F4 CRC64;

Query Match 23.2%; Score 22; DB 1; Length 16;
Best Local Similarity 57.1%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 DADQRII 8
I::: I:::
Db 10 DKDQELI 16

RESULT 10
HBB2_UROHA STANDARD; PRT; 19 AA.
ID HBB2_UROHA
AC P18992;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hemoglobin beta-2 chain (Fragment).
OS Uromastix hardwickii (Indian spiny-tailed lizard).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Uromastycinae;
OC Uromastix.
OX NCBI_TaxID=40250;
RN [1]
RP SEQUENCE.
RX MEDLINE=84029159; PubMed=6628672;
RA Naqvi S., Zaidi Z.H., von Bahr-Lindstroem H., Carlquist M.,

Joernvall H.;
RT "Characterization of hemoglobin from the lizard Uromastix
hardwickii".
RL FEBS Lett. 162:290-295(1983).
CC -|- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CC VARIOUS PERIPHERAL TISSUES.
CC -|- SUBUNIT: HETEROTETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -|- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -|- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR; A05305; A05305.
DR InterPro; IPR000971; Globin.
DR PROSITE; PS01033; GLOBIN; PARTIAL.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT NON_TER 1
FT NON_TER 19
SQ SEQUENCE 19 AA; 1914 MW; C40AD8EA30019057 CRC64;

Query Match 23.2%; Score 22; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 FGNI 18
I::: I:::
Db 5 FGNI 8

RESULT 11
CISP_TRIVA STANDARD; PRT; 22 AA.
ID CISP_TRIVA
AC P33404;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteine proteinase (EC 3.4.22.-) (Fragment).
OS Trichomonas vaginalis.
OC Eukaryota; Parabasalida; Trichomonadida; Trichomonadidae;
OC Trichomonadinae; Trichomonas.
OX NCBI_TaxID=5722;
RN [1]
RP SEQUENCE.
RX MEDLINE=93307628; PubMed=8319888;
RA Irvine J.W., Coombs G.H., North M.J.;
RT "Purification of cysteine proteinases from trichomonads using
baicitracin-Sepharose".
RL FEBS Microbiol. Lett. 110:113-120(1993).
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.
DR MEROPS; C01.082;
DR InterPro; IPR000169; SHprot_acsite.
DR PROSITE; PS00139; THIOLEPROTEASE_CYS; PARTIAL.
DR PROSITE; PS00639; THIOLEPROTEASE_HIS; PARTIAL.
DR PROSITE; PS00640; THIOLEPROTEASE_ASN; PARTIAL.
KW Hydrolase; Thiol protease.
FT UNSURE 1
FT NON_TER 22
SQ SEQUENCE 22 AA; 2398 MW; 0EE40FD86661ACCB CRC64;

Query Match 23.2%; Score 22; DB 1; Length 22;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DDADQR 6
I::: I:::
Db 3 DDSDWR 8

RESULT 12
NUO5_SOLTU STANDARD; PRT; 23 AA.
ID NUO5_SOLTU
AC P80262;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NADH-ubiquinone oxidoreductase 29 kDa subunit (EC 1.6.5.3)

CC -1- FUNCTION: INVOLVED IN THE BINDING OF FMET-TENA AND, HENCE, IN THE
CC INITIATION OF TRANSLATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE S13P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro: IPR001892; Ribosomal_S13.
DR Pfam: PF00416; Ribosomal_S13; 1.
DR PROSITE: PS00646; RIBOSOMAL_S13; PARTIAL.
KW Ribosomal protein.. 24
FT NON_TER 24
SQ SEQUENCE 24 AA; 2688 MW; 337F88EA0294D2A8 CRC64;
Query Match 23.2%; Score 22; DB 1; Length 24;
Best Local Similarity 30.8%; Pred. No. 1.6e+03;
Matches 4; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 4 DQRIKQLEYFG 16
Db 11 NKRVDVATYIG 23

Search completed: April 23, 2003, 13:43:52
Job time : 5.75169 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:36:43 ; Search time 9.40449 Seconds
(without alignments)
183.999 Million cell updates/sec

Title: US-09-836-073-17
Sequence: 1 DDADQRIKQLEYFGNI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 4984

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	33.7	25	2 A60286	heat-stable serine
2	30	31.6	18	2 PH1368	Ig heavy chain DJ
3	29	30.5	25	2 S67973	apolipoprotein Ca
4	28	29.5	15	2 PH0770	T-cell receptor be
5	28	29.5	19	2 B53145	high conductance c
6	28	29.5	25	2 CS7001	endo-1,4-beta-xyla
7	28	29.5	25	2 PC1314	large granule LI c
8	27	28.4	12	2 PH0771	T-cell receptor be
9	27	28.4	23	2 PH1722	Ig heavy chain V r
10	27	28.4	23	2 PH1727	Ig heavy chain V r
11	27	28.4	24	2 S56003	glucan 1,3-beta-gl
12	27	28.4	25	2 S56002	glucan 1,3-beta-gl
13	26	27.4	17	2 S57519	T cell receptor be
14	26	27.4	17	2 S57556	T cell receptor be
15	26	27.4	20	2 PN0133	pepsin (EC 3.4.23.
16	26	27.4	23	2 S65379	cytochrome-c oxida
17	26	27.4	24	2 PH1696	Ig heavy chain V r
18	25	26.3	9	2 S56004	glucan 1,3-beta-gl
19	25	26.3	22	2 PH1325	Ig heavy chain DJ
20	25	26.3	22	2 PH1679	Ig heavy chain V r
21	25	26.3	23	2 PH1694	Ig heavy chain V r
22	25	26.3	23	2 PH1707	Ig heavy chain V r
23	25	26.3	23	2 PH1725	Ig heavy chain V r
24	25	26.3	23	2 PH1723	Ig heavy chain V r
25	25	26.3	24	2 PT0258	Ig heavy chain CDR
26	25	26.3	24	2 PH1683	Ig heavy chain V r
27	25	26.3	24	2 PH1685	Ig heavy chain V r
28	25	26.3	24	2 PH1698	Ig heavy chain V r
29	25	26.3	24	2 PH1710	Ig heavy chain V r

30	25	26.3	24	2 PH1712	Ig heavy chain V r
31	25	26.3	24	2 PH1713	Ig heavy chain V r
32	25	26.3	25	2 PH1686	Ig heavy chain V r
33	25	26.3	25	2 PH1700	Ig heavy chain V r
34	25	26.3	25	2 PH1716	Ig heavy chain V r
35	24.5	25.8	25	2 PH1715	Ig heavy chain V r
36	24	25.3	12	2 S25056	Ig heavy chain - m
37	24	25.3	12	2 S57570	T cell receptor V-
38	24	25.3	14	2 S57569	T cell receptor V-
39	24	25.3	14	2 S57638	Ig heavy chain DJ
40	24	25.3	18	2 PH1323	Ig heavy chain CRD
41	24	25.3	19	2 PT0332	hypothetical prote
42	24	25.3	20	2 C49753	hypothetical prote
43	24	25.3	20	2 B30208	lambda I12 protein
44	24	25.3	20	2 PC2248	Ig heavy chain V r
45	24	25.3	21	2 PH1730	Ig heavy chain V r

ALIGNMENTS

RESULT 1

A60286
heat-stable serine proteinase (EC 3.4.21.-) - Thermomonospora fusca (strain YX) (N:Alternate names: IX-proteinase
C:Species: Thermomonospora fusca
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Dec-1996
C:Accession: A60286
R:Kristjansson, M.M.; Kinsella, J.E.
Int. J. Pept. Protein Res. 36, 201-207, 1990
A:Title: Heat stable proteinase from Thermomonospora fusca. Characterization as a
A:Reference number: A60286; MUID:91107200; PMID:2132918
A:Accession: A60286
A:Molecule type: protein
A:Residues: 1-25 <KRI>
C:Keywords: extracellular protein; hydrolase; serine proteinase

Query Match 33.7%; Score 32; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 YVFGN 17
|||||
DB 10 YVFGN 14

RESULT 2

PH1368
Ig heavy chain DJ region (clone C111-112) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PH1368
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor
A:Reference number: PH1302; MUID:93094761; PMID:1460419
A:Accession: PH1368
A:Molecule type: DNA
A:Residues: 1-18 <WAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin

Query Match 31.6%; Score 30; DB 2; Length 18;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 LEYFVG 16
:||||:
DB 5 MEYVYG 10

RESULT 3
S67973

C;Keywords: T-cell receptor

Query Match 28.4%; Score 27; DB 2; Length 12;
Best Local Similarity 71.4%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 QLEYFVG 16
| | | | |
Db 6 QYEQYFG 12

RESULT 9

PHI722

Ig heavy chain V region (clone GCC-2) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999

C;Accession: PH1722

R;McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.

J. Exp. Med. 178, 295-307, 1993

A;Title: Antigen-driven B cell differentiation in vivo.

A;Reference number: PH1675; MUID:93301607; PMID:8315385

A;Accession: PH1722

A;Molecule type: mRNA

A;Residues: 1-23 <MCH>

A;Experimental source: B cell

A;Note: the authors translated the codon ACA for residue 13 as Ala

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

Query Match 28.4%; Score 27; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 6.5e+02;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 10 QLEYFGNI 18
: | | | |
Db 12 ETRYYGSL 20

RESULT 10

PHI727

Ig heavy chain V region (clone GCC-7) - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999

C;Accession: PH1727

R;McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.

J. Exp. Med. 178, 295-307, 1993

A;Title: Antigen-driven B cell differentiation in vivo.

A;Reference number: PH1675; MUID:93301607; PMID:8315385

A;Accession: PH1727

A;Molecule type: mRNA

A;Residues: 1-23 <MCH>

A;Experimental source: B cell

A;Note: the authors translated the codon ACA for residue 13 as Ala

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

Query Match 28.4%; Score 27; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 6.5e+02;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 10 QLEYFGNI 18
: | | | |
Db 12 ETRYYGSL 20

RESULT 11

S56003

glucan 1,3-beta-glucosidase (EC 3.2.1.58) GNI1, extracellular - fungus (Acremonium persi

N;Alternate names: (1-3)-beta-D-glucan glucosylase GNI1

C;Species: Acremonium persicinum

C;Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 07-May-1999

C;Accession: S56003

R;Pitson, S.M.; Seviour, R.J.; McDougall, B.M.; Woodward, J.R.; Stone, B.A.

Biochem. J. 308, 733-741, 1995

A;Title: Purification and characterization of three extracellular (1->3)-beta-D-gl

A;Reference number: S56002; MUID:97104268; PMID:8948426

A;Accession: S56003

A;Molecule type: protein

A;Residues: 1-24 <PIT>

C;Keywords: glycosidase; hydrolase

Query Match 28.4%; Score 27; DB 2; Length 24;
Best Local Similarity 42.9%; Pred. No. 6.8e+02;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFEGNI 18
: | | | |
Db 3 DWFEGNV 9

RESULT 12

S56002

glucan 1,3-beta-glucosidase (EC 3.2.1.58) GNI1, extracellular - fungus (Acremonium

N;Alternate names: (1-3)-beta-D-glucan glucosylase GNI

C;Species: Acremonium persicinum

C;Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 07-May-1999

C;Accession: S56002

R;Pitson, S.M.; Seviour, R.J.; McDougall, B.M.; Woodward, J.R.; Stone, B.A.

Biochem. J. 308, 733-741, 1995

A;Title: Purification and characterization of three extracellular (1->3)-beta-D-gl

A;Reference number: S56002; MUID:97104268; PMID:8948426

A;Accession: S56002

A;Molecule type: protein

A;Residues: 1-25 <PIT>

C;Keywords: glycosidase; hydrolase

Query Match 28.4%; Score 27; DB 2; Length 25;
Best Local Similarity 42.9%; Pred. No. 7.1e+02;
Matches 3; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFEGNI 18
: | | | |
Db 3 DWFEGNV 9

RESULT 13

S57519

T cell receptor beta chain V region - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 10-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999

C;Accession: S57519

R;Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaeet, V.P.

submitted to the EMBL Data Library, June 1995

A;Description: T cell receptor repertoire for a viral epitope in humans is diversi

A;Reference number: S57494

A;Accession: S57519

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-17 <BUR>

A;Cross-references: EMBL:Z49930; NID:g887494; PIDN:CAA90176.1; PID:g887495

C;Keywords: T-cell receptor

Query Match 27.4%; Score 26; DB 2; Length 17;
Best Local Similarity 55.6%; Pred. No. 6.8e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 IKQLEYFVG 16
| | | | |
Db 9 ISSYEQVFG 17

RESULT 14

S57556

T cell receptor beta chain V-D-J region (clone PP7 and clone TF1) - human (fragmen

C;Species: Homo sapiens (man)

C;Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999

C;Accession: S57556; S57557
 R;Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Arguet, V.P.
 submitted to the EMBL Data Library, June 1995
 A;Description: T cell receptor repertoire for a viral epitope in humans is diversified
 C;Reference number: S57494
 A;Accession: S57556
 A;Molecule type: mRNA
 A;Residues: 1-17 <BUR>
 A;Cross-references: EMBL:Z49926; NID:g887498; PIDN:CAA90172.1; PID:g887499
 A;Experimental source: clone PP7
 A;Accession: S57557
 A;Molecule type: mRNA
 A;Residues: 1-17 <BUR>
 A;Cross-references: EMBL:Z49928; NID:g887502; PIDN:CAA90174.1; PID:g887503
 A;Experimental source: clone TP1
 C;Keywords: T-cell receptor

Query Match 27.4%; Score 26; DB 2; Length 17;
 Best Local Similarity 55.6%; Pred. No. 6.8e+02;
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 IKOLEYVFG 16
 | | | | |
 Db 9 ISSYEQYFG 17

RESULT 15
 PN0133
 pepsin (EC 3.4.23.-) 1 - horse (fragment)
 N;Alternate names: pepsin (pi 1.8)
 C;Species: Equus caballus (domestic horse)
 C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 07-May-1999
 C;Accession: PN0133
 R;Gonchar, M.V.; Lavrenova, G.I.; Rudenskaya, G.N.; Gaida, A.V.; Stepanov, V.M.
 Biochimia 49, 1026-1037, 1984
 A;Title: Multiple forms of horse pepsin.
 A;Reference number: PN0133; MUID:84281135; PMID:6432065
 A;Accession: PN0133
 A;Molecule type: protein
 A;Residues: 1-20 <GON>
 A;Note: article in Russian with English abstract
 C;Superfamily: pepsin
 C;Keywords: aspartic proteinase; hydrolase; protein digestion

Query Match 27.4%; Score 26; DB 2; Length 20;
 Best Local Similarity 71.4%; Pred. No. 8.2e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 12 EYFNGI 18
 | | | | |
 Db 12 EYFNGI 18

Search completed: April 23, 2003, 13:48:55
 Job time : 9.40449 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:43:09 ; Search time 9.91011 Seconds
(without alignments)
145.542 Million cell updates/sec

Title: US-09-836-073-17
Perfect score: 95
Sequence: 1 DDADQRIIKOLEYFGNI 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 61792

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	95	100.0	18	9	US-09-836-073-17
2	55	57.9	18	9	US-09-836-073-15
3	50	52.6	16	9	US-09-836-073-19
4	45	47.4	18	9	US-09-836-073-2
5	44	46.3	17	9	US-09-836-073-13
6	44	46.3	18	9	US-09-836-073-1
7	44	46.3	18	9	US-09-836-073-4
8	44	46.3	18	9	US-09-836-073-9
9	44	46.3	18	9	US-09-836-073-14
10	43	45.3	18	9	US-09-836-073-5
11	39	41.1	18	9	US-09-836-073-3
12	36	37.9	18	9	US-09-836-073-11
13	36	37.9	18	9	US-09-836-073-12
14	35	36.8	18	9	US-09-836-073-7
15	35	36.8	18	9	US-09-836-073-10
16	35	36.8	21	10	US-09-815-306-8
17	33.5	35.3	19	9	US-09-836-073-16
18	32.5	34.2	18	9	US-09-836-073-18
19	32	33.7	21	10	US-09-815-306-9

20	31	32.6	21	10	US-09-815-306-16	Sequence 16, Appl
21	30	31.6	17	9	US-10-011-588-2	Sequence 2, Appl
22	30	31.6	19	9	US-09-880-748-2928	Sequence 2928, Ap
23	29	30.5	24	9	US-09-991-548-3	Sequence 3, Appl
24	28.5	30.0	19	9	US-09-880-748-3113	Sequence 3113, Ap
25	28.5	30.0	21	10	US-09-764-017-2	Sequence 2, Appl
26	28.5	30.0	21	10	US-09-764-017-3	Sequence 3, Appl
27	28	29.5	16	9	US-09-880-748-3047	Sequence 3047, Ap
28	28	29.5	16	9	US-10-253-443-1	Sequence 1, Appl
29	28	29.5	16	10	US-09-835-124-1	Sequence 8, Appl
30	28	29.5	18	9	US-09-836-073-8	Sequence 11, Appl
31	28	29.5	19	9	US-09-880-748-2987	Sequence 2987, Ap
32	28	29.5	21	9	US-09-880-748-3010	Sequence 3010, Ap
33	28	29.5	25	8	US-08-424-550B-386	Sequence 386, App
34	27	28.4	12	9	US-09-977-797A-20	Sequence 20, Appl
35	27	28.4	13	9	US-09-880-748-3091	Sequence 3091, Ap
36	27	28.4	18	9	US-09-880-748-3060	Sequence 3060, Ap
37	27	28.4	18	9	US-10-084-813-138	Sequence 138, App
38	27	28.4	18	9	US-10-084-813-139	Sequence 139, App
39	27	28.4	18	9	US-10-084-813-140	Sequence 140, App
40	27	28.4	18	9	US-10-084-813-141	Sequence 141, App
41	27	28.4	19	9	US-10-214-188-21	Sequence 21, Appl
42	27	28.4	19	9	US-09-880-748-3038	Sequence 3038, Ap
43	27	28.4	22	9	US-09-984-271-218	Sequence 218, App
44	26.5	27.9	16	9	US-09-998-425-8	Sequence 8, Appl
45	26.5	27.9	16	9	US-09-997-977-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-09-836-073-17
; Sequence 17, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; TYPE: PRT
; LENGTH: 18
; ORGANISM: C. elegans
US-09-836-073-17

Query Match 100.0%; Score 95; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DDADQRIIKOLEYFGNI 18
Db 1 DDADQRIIKOLEYFGNI 18

RESULT 2
US-09-836-073-15
; Sequence 15, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match 46.3%; Score 44; DB 9; Length 18;
Best Local Similarity 50.0%; Pred. No. 0.79;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 DQRIIKOLEYFNG 17
Db 4 EAKICHOIEYFNG 17

RESULT 8.

US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match 46.3%; Score 44; DB 9; Length 18;
Best Local Similarity 50.0%; Pred. No. 0.79;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 DQRIIKOLEYFNG 17
Db 4 EAKICHOIEYFNG 17

RESULT 9

US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14

Query Match 46.3%; Score 44; DB 9; Length 18;
Best Local Similarity 50.0%; Pred. No. 0.79;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 DQRIIKOLEYFNG 17
Db 4 EAKICHOIEYFNG 17

RESULT 10

US-09-836-073-5
; Sequence 5, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-5

Query Match 45.3%; Score 43; DB 9; Length 18;
Best Local Similarity 50.0%; Pred. No. 1.2;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 DQRIIKOLEYFNG 17
Db 4 EAKICHOIEYFNG 17

RESULT 11

US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match 41.1%; Score 39; DB 9; Length 18;
Best Local Similarity 54.5%; Pred. No. 5.1;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 6 RIHKQLEYFG 16
: | | : | | | |
Db 6 KICHQIQYFG 16

RESULT 12

US-09-836-073-11

; Sequence 11, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836.073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 11

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-11

Query Match 37.9%; Score 36; DB 9; Length 18;

Best Local Similarity 42.9%; Pred. No. 15;

Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 4 DORIKQLEYFGN 17
: | | | : | | | |
Db 4 EAKICHQIEYQFGD 17

RESULT 13

US-09-836-073-12

; Sequence 12, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836.073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-12

Query Match

Best Local Similarity 37.9%; Score 36; DB 9; Length 18;

Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 4 DORIKQLEYFGN 17
: | | | : | | | |
Db 4 EAKICHQIEYQFGD 17

RESULT 14

US-09-836-073-7

; Sequence 7, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836.073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-7

Query Match 36.8%; Score 35; DB 9; Length 18;

Best Local Similarity 42.9%; Pred. No. 22;

Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 4 DORIKQLEYFGN 17
: | | | : | | | |
Db 4 EAKICHQIEYQGD 17

RESULT 15

US-09-836-073-10

; Sequence 10, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836.073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-10

Query Match

Best Local Similarity 36.8%; Score 35; DB 9; Length 18;

Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 4 DORIKQLEYFGN 17
: | | | : | | | |
Db 4 EAKICHQIEYQGD 17

Search completed: April 23, 2003, 13:52:11

Job time : 10.9101 secs

[illegible]

COUNTRY: USA

```

;
;
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,894B
; FILING DATE: No. 5965699ember 6, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid sequence
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-743-894B-1

```

Query Match 31.6%; Score 30; DB 2; Length 17;
 Best Local Similarity 54.5%; Pred. No. 78;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 DDADQRIIKQL 11
DB 7 DEANQRATKML 17

```

RESULT 6

```

US-08-743-894B-19
; Sequence 19, Application US/08743894B
; Patent No. 5965699
; GENERAL INFORMATION:
; APPLICANT: James J. Schmidt
; APPLICANT: Karen A. Bostian
; TITLE OF INVENTION: Assay for the Proteolytic Activity of Serotype A Neurotoxin fr
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MCMR-JA Attn:John Moran-Patent Atty
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,894B
; FILING DATE: No. 5965699ember 6, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 19:

```

```

;
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid sequence
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; FEATURE:
; LOCATION: 4
; OTHER INFORMATION: Xaa represents Abu, or 2-Aminobutyric Acid
; US-08-743-894B-19

```

Query Match 31.6%; Score 30; DB 2; Length 17;
 Best Local Similarity 54.5%; Pred. No. 78;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 DDADQRIIKQL 11
DB 7 DEANQRATKML 17

```

RESULT 7

```

US-08-743-894B-25
; Sequence 25, Application US/08743894B
; Patent No. 5965699
; GENERAL INFORMATION:
; APPLICANT: James J. Schmidt
; APPLICANT: Karen A. Bostian
; TITLE OF INVENTION: Assay for the Proteolytic Activity of Serotype A Neuroto:
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MCMR-JA Attn:John Moran-Patent Atty
; STREET: USA MRC - 504 Scott Street
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,894B
; FILING DATE: No. 5965699ember 6, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid sequence
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; FEATURE:
; LOCATION: 14
; OTHER INFORMATION: Xaa represents Abu, or 2-Aminobutyric Acid
; US-08-743-894B-25

```

Query Match 31.6%; Score 30; DB 2; Length 17;
 Best Local Similarity 54.5%; Pred. No. 78;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 DDADQRIIKQL 11
DB 7 DEANQRATKML 17

```

RESULT 8

US-08-743-894B-26
; Sequence 26, Application US/08743894B
; Patent No. 5965699
; GENERAL INFORMATION:
; APPLICANT: James J. Schmidt
; APPLICANT: Karen A. Bostian
; TITLE OF INVENTION: Assay for the Proteolytic Activity of Serotype A Neurotoxin fr
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MCMR-JA Attn:John Moran-Patent Atty
; STREET: USA MRMC - 504 Scott Street
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08743,894B
; FILING DATE: No. 5965699ember 6, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; INFORMATION FOR SEQ ID NO: 26:
; LENGTH: 17 amino acids
; TYPE: amino acid sequence
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; FEATURE:
US-08-743-894B-26

Query Match 31.6%; Score 30; DB 2; Length 17;
Best Local Similarity 54.5%; Pred. No. 78;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDADQRIIKQL 11
|:|:|:|
Db 7 DEANQRTKML 17

RESULT 9

US-08-743-894B-27
; Sequence 27, Application US/08743894B
; Patent No. 5965699
; GENERAL INFORMATION:
; APPLICANT: James J. Schmidt
; APPLICANT: Karen A. Bostian
; TITLE OF INVENTION: Assay for the Proteolytic Activity of Serotype A Neurotoxin fr
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MCMR-JA Attn:John Moran-Patent Atty
; STREET: USA MRMC - 504 Scott Street
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh

; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08743,894B
; FILING DATE: No. 5965699ember 6, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; INFORMATION FOR SEQ ID NO: 27:
; LENGTH: 17 amino acids
; TYPE: amino acid sequence
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; FEATURE:
; LOCATION: 13
; OTHER INFORMATION: Xaa represent Abu, or 2-Aminobutyric Acid
US-08-743-894B-27

Query Match 31.6%; Score 30; DB 2; Length 17;
Best Local Similarity 54.5%; Pred. No. 78;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDADQRIIKQL 11
|:|:|:|
Db 7 DEANQRTKML 17

RESULT 10

US-08-743-894B-36
; Sequence 36, Application US/08743894B
; Patent No. 5965699
; GENERAL INFORMATION:
; APPLICANT: James J. Schmidt
; APPLICANT: Karen A. Bostian
; TITLE OF INVENTION: Assay for the Proteolytic Activity of Serotype A Neurotox
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MCMR-JA Attn:John Moran-Patent Atty
; STREET: USA MRMC - 504 Scott Street
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08743,894B
; FILING DATE: No. 5965699ember 6, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:

LENGTH: 17 amino acids
 TYPE: amino acid sequence
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 FEATURE:
 US-08-743-894B-36

Query Match 31.6%; Score 30; DB 2; Length 17;
 Best Local Similarity 54.5%; Pred. No. 78;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDADQRIIKQL 11
 I:|:|:|:|
 Db 7 DEANQRATKML 17

RESULT 11
 US-08-743-894B-37
 ; Sequence 37, Application US/08743894B
 ; Patent No. 5965699

GENERAL INFORMATION:
 APPLICANT: James J. Schmidt
 TITLE OF INVENTION: Assay for the Proteolytic Activity of Serotype A Neurotoxin fr
 NUMBER OF SEQUENCES: 56
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MCMR-JA Attn:John Moran-Patent Atty
 STREET: USA MRC - 504 Scott Street
 CITY: FORT DETRICK
 STATE: MARYLAND
 COUNTRY: USA
 ZIP: 21702-5012
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Macintosh 7.0
 SOFTWARE: Microsoft Word 6.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/743,894B
 FILING DATE: No. 5965699ember 6, 1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Charles H. Harris
 REGISTRATION NUMBER: 34,616
 REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 619-2065
 TELEFAX: (301) 619-7714
 INFORMATION FOR SEQ ID NO: 37:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 17 amino acids
 TYPE: amino acid sequence
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 FEATURE:

US-08-743-894B-37
 Query Match 31.6%; Score 30; DB 2; Length 17;
 Best Local Similarity 54.5%; Pred. No. 78;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDADQRIIKQL 11
 I:|:|:|:|
 Db 7 DEANQRATKML 17

RESULT 12
 US-08-743-894B-39
 ; Sequence 39, Application US/08743894B
 ; Patent No. 5965699

GENERAL INFORMATION:
 APPLICANT: James J. Schmidt
 TITLE OF INVENTION: Assay for the Proteolytic Activity of Serotype A Neuroto
 NUMBER OF SEQUENCES: 56
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MCMR-JA Attn:John Moran-Patent Atty
 STREET: USA MRC - 504 Scott Street
 CITY: FORT DETRICK
 STATE: MARYLAND
 COUNTRY: USA
 ZIP: 21702-5012
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Macintosh 7.0
 SOFTWARE: Microsoft Word 6.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/743,894B
 FILING DATE: No. 5965699ember 6, 1996
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Charles H. Harris
 REGISTRATION NUMBER: 34,616
 REFERENCE/DOCKET NUMBER:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 619-2065
 TELEFAX: (301) 619-7714
 INFORMATION FOR SEQ ID NO: 39:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 17 amino acids
 TYPE: amino acid sequence
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 FEATURE:
 US-08-743-894B-39

Query Match 31.6%; Score 30; DB 2; Length 17;
 Best Local Similarity 54.5%; Pred. No. 78;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDADQRIIKQL 11
 I:|:|:|:|
 Db 7 DEANQRATKML 17

RESULT 13
 US-08-819-286-9
 ; Sequence 9, Application US/08819286
 ; Patent No. 6169074
 GENERAL INFORMATION:
 APPLICANT: Montal, Mauricio
 TITLE OF INVENTION: PEPTIDE INHIBITORS OF
 TITLE OF INVENTION: NEUROTRANSMITTER SECRETION BY NEURONAL CELLS
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 4225 Executive Square, Suite 1400
 CITY: La Jolla
 STATE: CA
 COUNTRY: USA
 ZIP: 92037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/819,286
 FILING DATE:

```

; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,599
; FILING DATE: 18-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Taylor, Stacy L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: 07349/005001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-819-286-9

```

```

Query Match 31.6%; Score 30; DB 4; Length 20;
Best Local Similarity 54.5%; Pred. No. 94;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 DDADQRIIKQL 11
   ||:|| |
Db 7 DEANQATKML 17

```

```

RESULT 14
US-08-408-604A-69
; Sequence 69, Application US/08408604A
; Patent No. 5801149
; GENERAL INFORMATION:
; APPLICANT: Shoelson, Steven
; TITLE OF INVENTION: INHIBITION OF SIGNAL TRANSDUCTION MOLECULES
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/408,604A
; FILING DATE: 21-MAR-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/134,558
; FILING DATE: 08-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/959,949
; FILING DATE: 09-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/722,359
; FILING DATE: 19-JUNE-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: JDP-014CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-408-604A-69
Query Match 30.5%; Score 29; DB 1; Length 13;
Best Local Similarity 38.5%; Pred. No. 85;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy 2 DADQRIIKQLEY 14
   ||:|| |
Db 1 DADEXLIPQGGFF 13

```

```

RESULT 15
US-08-743-894B-24
; Sequence 24, Application US/08743894B
; Patent No. 5965699
; GENERAL INFORMATION:
; APPLICANT: James J. Schmidt
; APPLICANT: Karen A. Bostian
; TITLE OF INVENTION: Assay for the Proteolytic Activity of Serotype A Neurotoxin
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MCMR-JA Attn: John Moran-Patent Atty
; STREET: USA MCMC - 504 Scott Street
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/743,894B
; FILING DATE: No. 5965699ember 6, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid sequence
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; FEATURE:
; LOCATION: 16
; OTHER INFORMATION: Xaa represents norleucine
; US-08-743-894B-24

```

```

Query Match 30.5%; Score 29; DB 2; Length 17;
Best Local Similarity 54.5%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 DDADQRIIKQL 11
   ||:|| |
Db 7 DEANQATKML 17

```

Search completed: April 23, 2003, 13:50:28
Job time : 9.49438 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:28:18 ; Search time 27.2022 seconds

(without alignments)
88.173 Million cell updates/sec

Title: US-09-836-073-17

Perfect score: 95

Sequence: 1 DDADQRIIKQLEYFNGNI 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 320064

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	44	46.3	18	21	AAV52200 Human la autoantigen
2	41	43.2	21	14	AA198394 La/SSB epitope 17..
3	33	34.7	14	19	AAW76974 Fusion immunoglobulin
4	33	34.7	14	22	AAW99328 Vaccine related MH
5	33	34.7	17	20	AAW44062 Human SNAP25 (amin
6	32	33.7	14	22	AAW97641 Human peptide #916
7	32	33.7	19	22	AU00696 Fugu neurofibromat
8	31	32.6	16	23	AU97828 Glucose dehydrogen
9	31	32.6	21	11	AAW07109 Antibody for detec
10	31	32.6	21	22	AAW92191 Signal transductio

11	30	31.6	12	20	AAV44061 Human SNAP25 (amin
12	30	31.6	15	23	AAU81241 Murine trkC antibo
13	30	31.6	17	20	AAU44021 Amino acids 187-20
14	30	31.6	17	20	AAV44039 Human SNAP25 (amin
15	30	31.6	17	20	AAV44045 Human SNAP25 (amin
16	30	31.6	17	20	AAV44046 Human SNAP25 (amin
17	30	31.6	17	20	AAV44047 Human SNAP25 (amin
18	30	31.6	17	20	AAV44056 Human SNAP25 (amin
19	30	31.6	17	20	AAV44057 Human SNAP25 (amin
20	30	31.6	17	20	AAV44059 Human SNAP25 (amin
21	30	31.6	17	23	ABG69065 Human polypeptide
22	30	31.6	18	19	AAW69133 Neuronal NOS bindi
23	30	31.6	19	22	AB15586 Human SNAP-25 N-te
24	30	31.6	19	23	ABP46917 Human Blys binding
25	30	31.6	20	18	AAW30100 Neurotransmitter s
26	29.5	31.1	15	22	ABW09465 Haemagglutination
27	29	30.5	10	22	ABG95217 Human complementar
28	29	30.5	11	22	ABP22066 HIV A03 motif vif
29	29	30.5	12	22	ABP18371 HIV B58 super moti
30	29	30.5	12	22	AAU02675 CD8 region of anti
31	29	30.5	13	19	AAW78582 SH2 domain binding
32	29	30.5	14	21	ABW51868 Human kinesin ligh
33	29	30.5	15	23	ABW08010 Human SNAP25 (amin
34	29	30.5	17	20	AAV44044 Human SNAP25 (amin
35	29	30.5	17	20	AAV44048 Human SNAP25 (amin
36	29	30.5	17	20	AAV44052 Human SNAP25 (amin
37	29	30.5	21	11	AAW05023 Papilloma virus ty
38	29	30.5	24	18	AAW32070 Insulin internallis
39	29	30.5	24	20	AAW39403 Human insulin rece
40	29	30.5	24	23	AAU78003 Human insulin rece
41	29	30.5	24	23	AAU83372 Antiviral composi
42	28.5	30.0	13	23	AAU83378 Antiviral composi
43	28.5	30.0	14	23	AAU83379 Antiviral composi
44	28.5	30.0	14	23	AAU83379 Antiviral composi
45	28.5	30.0	14	23	AAU83382 Antiviral composi

ALIGNMENTS

RESULT 1

AAV52200
ID AAV52200 standard; peptide; 18 AA.

XX AC AAV52200;

XX DT 14-MAR-2000 (first entry)

XX DE Human la autoantigen peptide (LAP).

XX KW La autoantigen; LAP; Internal ribosome entry site; IRES; translation;
KW viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
KW coronavirus; hepatitis virus; rhabdovirus; adenovirus; coxsackie virus;
KW parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
KW vesicular stomatitis virus.

OS Homo sapiens.

XX PN WO9961613-A2.

XX PD 02-DEC-1999.

XX PF 21-MAY-1999; 99WO-US11281.

XX PR 22-MAY-1998; 98US-0086527.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Das S, Dasgupta A;

XX DR WPI; 2000-062712/05.

XX

CC human, IgH chain fused in frame at its N-terminus to one or more human
 CC immune deficiency virus (HIV) gp120 epitopes. Fusion Ig proteins and/or
 CC transfected cells are used to tolerate subjects to gp120 epitopes and to
 CC maintain this tolerance, particularly for treatment of HIV infection,
 CC optionally together with other therapeutic/prophylactic agents such as
 CC vaccines, chemotherapeutic agents and immune response modifiers. Such
 CC proteins can be used against other diseases where an immune response is
 CC deleterious, e.g. microbial infection, tumours or autoimmune disease.
 CC Induction of tolerance suppresses production of antibodies against gp120,
 CC so prevents or inhibits 'bystander' apoptosis of uninfected T cells that
 CC are bound to gp120 protein, maximising induction of protective antiviral
 CC T cell immunity.

XX Sequence 14 AA;

Query Match 34.7%; Score 33; DB 19; Length 14;
 Best Local Similarity 50.0%; Pred. No. 66;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 RIIRKOLEYYFGN 17
 :|:|:| |||
 Db 1 QIVKKLRQFGN 12

RESULT 4

AAAM99328
 ID AAM99328 standard; Peptide; 14 AA.

XX AC AAM99328;

XX DT 07-DEC-2001 (first entry)

DE DE Vaccine related MHC ligand peptide SEQ ID NO:431.

XX KW Glutamic acid; glutamine; vaccine; major histocompatibility complex;
 KW MHC; immunomodulator; antiallergic; endocrine; neuroprotectant;
 KW virucidal; bactericidal; antiparasitic; fungicidal; cytostatic;
 KW medicine; pharmaceutical; immune disorder; immune deficiency;
 KW autoimmune; hypersensitivity; allergy; graft rejection; infection;
 KW hormonal disorder; central nervous system disease; cancer; melanoma;
 KW anti-melanoma vaccine; human immunodeficiency virus.

OS Human immunodeficiency virus type 1.

XX PN WO200170772-A2.

XX PD 27-SEP-2001.

XX PF 22-MAR-2001; 2001WO-FR00872..

XX PR 23-MAR-2000; 2000FR-0003711.

XX PA (FABR) FABRE MEDICAMENT SA PIERRE.

XX PI Klinguer-Hamouir C, Corvaia N, Beck A, Goetsch L;

XX DR WPI; 2001-611470/70.

XX PT Stabilized pharmaceutical containing N-terminal glutamic acid or
 PT glutamine, useful e.g. in anti-melanoma vaccines, is an addition salt
 PT with strong acid

XX PS Claim 9; Page 105; 149pp; French.

XX CC The present invention describes a pharmaceutical compound (I) that
 CC contains an N-terminal glutamic acid (Glu) or glutamine (Gln) residue
 CC in the form of an addition salt with a strong, physiologically
 CC acceptable acid (II). Also described are: (a) a pharmaceutical
 CC composition containing at least one (I); (b) a vaccine containing at
 CC least one (I) where this is a major histocompatibility complex (MHC)
 CC ligand (Ia); (c) a method for in vitro diagnosis of diseases associated
 CC with the presence of (Ia); (d) a kit for method (c) that includes a (Ia);
 CC and (e) a process for preparing (I). (I) has immunomodulator, endocrine,

CC antiallergic, neuroprotectant, virucidal, bactericidal, antiparasitic,
 CC fungicidal and cytostatic activities. (I) are useful, in human or
 CC veterinary medicine, in pharmaceutical compositions (for treating immune
 CC disorders, e.g. immune deficiency, autoimmune states, hypersensitivity,
 CC allergy, graft rejection, infection, hormonal disorders and central
 CC nervous system diseases), also, where (I) is a MHC ligand (Ia), in
 CC vaccines for treatment or prevention of: (i) viral, bacterial, parasitic
 CC or fungal infections; or (ii) of cancers. A particular application is in
 CC anti-melanoma vaccines. (I) are also useful for in vitro diagnosis of
 CC diseases associated with interactions between MHC and (I), e.g. melanoma
 CC and human immunodeficiency virus infection. AAM98898 to AAM99592
 CC represent peptides which can be used in pharmaceutical compounds from
 CC the present invention.

XX Sequence 14 AA;

Query Match 34.7%; Score 33; DB 22; Length 14;
 Best Local Similarity 50.0%; Pred. No. 66;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 RIIRKOLEYYFGN 17
 :|:|:| |||
 Db 1 QIVKKLRQFGN 12

RESULT 5

AAAY44062

ID AAY44062 standard; peptide; 17 AA.

XX AC AAY44062;

XX DT 18-JAN-2000 (first entry)

DE DE Human SNAP25 (amino acids 187-203) analogue A13L.

XX KW Enzymatic assay; quantitation; type A botulinum neurotoxin; proteolysis;
 KW fluorescamine; detection; human; synaptosomal protein; SNAP25;
 KW hydrolysis; amino group.

OS Homo sapiens.

OS Synthetic.

XX PN US5965699-A.

XX PD 12-OCT-1999.

XX PF 06-NOV-1996; 96US-0743894.

XX PR 06-NOV-1996; 96US-0743894.

XX PA (USSA) US SEC OF ARMY.

XX PI Bostian KA, Schmidt JJ;

XX DR WPI; 1999-579939/49.

XX PT Quantitation of type A botulinum toxin -

XX PS Disclosure; Column 9-10; 28pp; English.

XX CC The invention relates to an enzymatic assay for the quantitation of
 CC type A botulinum toxin, by determining the proteolytic activity of
 CC botulinum neurotoxin type A using fluorescamine detection. Botulinum
 CC toxin A has been shown to cleave the synaptosomal neurotransmitter
 CC peptide SNAP25 between residues 197-198. The method comprises adding
 CC an analogue (e.g. AAY44022-Y44076) of the SNAP25 peptide (AAY44021,
 CC amino acids 187-203 of human SNAP25) to a sample containing the
 CC botulinum toxin A so that hydrolysis of the peptide is initiated, then
 CC stopping hydrolysis of the peptide at different time points; and
 CC measuring the amount of hydrolysis at each time point by combining with a
 CC label capable of detecting free amino groups resulting from the
 CC hydrolysis. The amount of botulinum toxin A present in the sample is
 CC determined by comparing measurements with the amount of label produced

CC from a known concentration of toxin measured under similar conditions.
 CC The method is useful for the quantitation of type A botulinum toxin.

SQ Sequence 17 AA;
 Query Match 34.7%; Score 33; DB 20; Length 17;
 Best Local Similarity 54.5%; Pred. No. 83;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 DDADQRIKQL 11
 I::I:: I I
 Db 7 DEANQRLTKML 17

RESULT 6
 AAM97641
 ID AAM97641 standard; Peptide; 14 AA.
 XX AC
 XX AAM97641;
 DT 24-JAN-2002 (first entry)
 XX
 DE Human peptide #916 encoded by a SNP oligonucleotide.
 XX
 KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
 KW neuroprotective; antimicrobial; gene therapy; vaccine; cancer;
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;
 KW cyclin; polymerase; histone; kinase; colony stimulating factor;
 KW complement related protein; cytochrome; kinesin; cytokine; interferon;
 KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
 KW multifactorial disease; autoimmune disease; infection;
 KW nervous system disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200147944-A2.
 XX
 PD 05-JUL-2001.
 XX
 PF 28-DEC-2000; 2000WO-US35498.
 XX
 PR 28-DEC-1999; 99US-0173419.
 PR 27-DEC-2000; 2000US-0173419.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI; 2001-465210/50.
 XX
 PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
 PT oncogenes and histones, useful for diagnosing and treating, e.g.
 PT cancer, autoimmune diseases and infections.
 XX
 PS Disclosure; Page 3868; 4143pp; English.
 XX
 CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
 CC encoding polymorphic variants of proteins related to amylases, amyloid
 CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
 CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
 CC complement related proteins, cytochromes, kinesins, cytokines,
 CC interferons, interleukins, G-protein coupled receptors and thioesterases.
 CC The present sequence is a peptide encoded by one such oligonucleotide.
 CC The oligonucleotides and the peptides encoded by them may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate expression of the proteins listed above. Disorders that may
 CC be prevented, diagnosed and/or treated include multifactorial diseases
 CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
 CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus
 CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
 CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
 CC system and an infection of pathogenic organisms.

SQ Sequence 14 AA;
 Query Match 33.7%; Score 32; DB 22; Length 14;
 Best Local Similarity 50.0%; Pred. No. 97;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 4 DQRIKQLEYFGN 17
 I::I:: I I I I
 Db 1 DQKGKREFERYGGN 14
 RESULT 7
 AAU00696
 ID AAU00696 standard; Peptide; 19 AA.
 XX AC
 XX AAU00696;
 DT 07-SEP-2001 (first entry)
 XX
 DE Fugu neurofibromatosis (NF1) protein fragment.
 XX
 KW Neurofibromatosis type 1; NF1; peripheral blood lymphocyte; PBL; EBV;
 KW Epstein-Barr virus; B-lymphoblastoid cell; phytohaemagglutinin; PHA;
 KW frame shift mutation; mis-sense mutation; silent mutation.
 XX
 OS Fugu sp.
 XX
 PN WO200129251-A2.
 XX
 PD 26-APR-2001.
 XX
 PF 18-OCT-2000; 2000WO-EP10255.
 XX
 PR 18-OCT-1999; 99EP-0870216.
 PR 05-JUN-2000; 2000EP-0870122.
 PR 16-JUN-2000; 2000US-0211629.
 XX
 PA (UYGE-) UNIV GENT.
 XX
 PI Messiaen L, Callens T;
 XX
 DR WPI; 2001-300341/31.
 XX
 PT Mutation analysis of NF1 gene by treating EBV transformed
 PT lymphoblastoid cell lines formed with lymphocytes of patient with
 PT protein synthesis inhibitor, and obtaining peptides by translating
 PT amplified RNA from cell line.
 XX
 PS Disclosure; Fig 16; 102pp; English.

The DNA sequences represent neurofibromatosis type 1 (NF1) peptide fragments. A method for mutation analysis of the NF1 gene involves isolating peripheral blood lymphocytes (PBL) of a patient, establishing Epstein-Barr virus (EBV) transformed B-lymphoblastoid cell line with isolated PBL, or short-term culturing of PBL by phytohaemagglutinin (PHA) stimulation, treating the cell line or short-term culture with protein synthesis inhibitor and immediately extracting RNA from the cultures. The RNA is then amplified and peptide fragments are obtained by in vitro transcription/translation of amplified fragments. Mutation analysis of NF1 is used for detection of frame shift, mis-sense and silent mutations in various exons of the gene. This is useful in screening for NF1 mutations in young children who are often oligosymptomatic. Efficacy of a drug or agent can be identified by a screening process in which the modulator is monitored in vitro using cell systems in which the defective NF1 gene is expressed. The sequences can be used to design drugs which modulate NF1 activity, by using knowledge of the structure of the NF1 protein and of specific defects of the various NF1 mutant proteins. The method allows for reliable analysis of mutations that are difficult to detect due to unstable or wrong-spliced transcripts.

SQ Sequence 19 AA;

Query Match 33.7%; Score 32; DB 22; Length 19;

Best Local Similarity 75.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DDADQRII 8
II IIII;
DB 5 DDFDQRII 12

RESULT 8
AAU97828
ID AAU97828 standard; Peptide; 16 AA.
XX
AC AAU97828;
XX
DT 27-AUG-2002 (first entry)
XX
DE Glucose dehydrogenase associated peptide.
XX
KW Glucose dehydrogenase; electrode; glucose sensor; glucose-assay;
XX substrate specificity; heat stability.
XX
OS Burkholderia cepacia.
XX
PN WO200236779-A1.
XX
PD 10-MAY-2002.
XX
PF 31-OCT-2001; 2001WO-JP09556.
XX
PR 31-OCT-2000; 2000JP-0332085.
XX
PR 24-NOV-2000; 2000JP-0357102.
XX
PR 12-SEP-2001; 2001JP-0276832.
XX
PA (SODE/) SODE K.
XX
PI Sode K;
XX
DR WPI; 2002-463413/49.
XX
PT Production of Burkholderia glucose dehydrogenase for use in glucose
PT sensor electrodes and glucose-assay kits in medicine, science and
PT industry
XX
PS Claim 12; Page 58; 61pp; Japanese.
XX
CC The invention describes a method of producing a glucose dehydrogenase
CC comprising culturing a Burkholderia microorganism and collecting the
CC product from the medium and/or the microbial cells. Glucose dehydrogenase
CC is useful in electrodes of glucose sensors and glucose-assay kits for
CC medicine, science and industry. Glucose dehydrogenase is economically
CC produced with high substrate specificity and improved heat stability to
CC provide long-term accuracy. This is the amino acid sequence of a glucose
CC dehydrogenase associated peptide described in the invention.
XX
SQ Sequence 16 AA;
Query Match 32.6%; Score 31; DB 23; Length 16;
Best Local Similarity 46.2%; Pred. No. 1.7e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 DDADQRII 13
II IIII;
DB 2 DAADPALVKRGEY 14

RESULT 9
AAU97109
ID AAU97109 standard; protein; 21 AA.
XX
AC AAU97109;
XX
DT 23-JAN-1991 (first entry)
XX

DE Antibody for detecting retinoblastoma.
XX
KW Retinoblastoma; antibody; tumours.
XX
PN EP390530-A.
XX
PD 03-OCT-1990.
XX
PF 28-MAR-1990; 90EP-0303297.
XX
PR 31-MAR-1989; 89US-0332082.
XX
PA (REDE-) RES DEV FOUNDATION.
XX
PI Fung YK, T'Ang A;
XX
WPI; 1990-299435/40.
XX
PT Antibodies to retinoblastoma gene prod. - used to detect RB1
PT protein expression to assess the stage and grade of tumours
XX
PS Claim 5; page 10; 19pp; English.
XX
CC This antibody (Ab) has specific binding affinity for the RB1 protein
CC prod. of the retinoblastoma (RB) gene. It has a sequence corresp.
CC to the C-terminal of the RB1 protein. It can be used for the
CC detection of RB1 protein expression to assess the stage and grade of
CC a wide variety of tumours for diagnosis and prognosis. It is useful
CC in the treatment of e.g. osteosarcomas, fibrosarcomas, carcinomas
CC of breast, ovary, bladder, lung, cervix), and soft tissue sarcomas
CC as well as RB's. Subpopulations of cells deficient in RB1 protein
CC expression can be identified. See also AAR07108 and AAR07110-11.
XX
SQ Sequence 21 AA;
Query Match 32.6%; Score 31; DB 11; Length 21;
Best Local Similarity 63.6%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 DADQRII 12
I II IIII
DB 9 DRSARIKQLE 19

RESULT 10
AAB92191
ID AAB92191 standard; Peptide; 21 AA.
XX
AC AAB92191;
XX
DT 22-JUN-2001 (first entry).
XX
DE Signal transduction reagents peptide SEQ ID NO:1367.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimide; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
XX
PR 10-SEP-1999; 99US-0153406.
XX
PR 15-OCT-1999; 99US-0159783.
XX
PA (CONJ-) CONJUCHEM INC.
XX

PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
 XX WPI; 2001-112059/12.
 XX
 PT Modifying and attaching therapeutic peptides to albumin prevents
 PT peptidase degradation, useful for increasing length of in vivo activity
 PT
 XX
 PS Disclosure; Page 643; 733pp; English.
 XX
 CC The present invention describes a modified therapeutic peptide (I)
 CC comprising a therapeutically active amino acid region (III) and a
 CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
 CC a less therapeutically active amino acid region (IV), which covalently
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a
 CC peptide stabilised therapeutic peptide composed of 3-50 amino acids.
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity
 CC in vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or
 CC reduces the action of peptidases to increase length of activity (half
 CC life) and specificity as bonding to large molecules decreases
 CC intracellular uptake and interference with physiological processes.
 CC AAB90829 to AAB92441 represent peptides which can be used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 21 AA;
 Query Match 32.6%; Score 31; DB 22; Length 21;
 Best Local Similarity 54.5%; Pred. NO. 2.3e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 OY 8 IKLEYFEGNI 18
 I: I:||||
 DB 7 IQAEWYFGKI 17
 RESULT 11
 AAY44061
 ID AAY44061 standard; peptide; 12 AA.
 XX
 AC AAY44061;
 XX
 DT 18-JAN-2000 (first entry)
 XX
 DE Human SNAP25 (amino acids 187-203) analogue #40.
 XX
 KW Enzymatic assay; quantitation; type A botulinum neurotoxin; proteolysis;
 KW fluorescamine; detection; human; synaptosomal protein; SNAP25;
 KW hydrolysis; amino group.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN US5965699-A.
 XX
 PD 12-OCT-1999.
 XX
 PF 06-NOV-1996; 96US-0743894.
 XX
 PR 06-NOV-1996; 96US-0743894.
 XX
 PA (USSA) US SEC OF ARMY.
 XX
 PI Bostian KA, Schmidt JU;
 XX
 DR WPI; 1999-579939/49.
 XX
 PT Quantitation of type A botulinum toxin -
 XX Disclosure; Column 9-10; 28pp; English.
 PS

XX The invention relates to an enzymatic assay for the quantitation of
 CC type A botulinum toxin, by determining the proteolytic activity of
 CC botulinum neurotoxin type A using fluorescamine detection. Botulinum
 CC toxin A has been shown to cleave the synaptosomal neurotransmitter
 CC peptide SNAP25 between residues 197-198. The method comprises adding
 CC an analogue (e.g. AAY44022-Y44076) of the SNAP25 peptide (AAY44021,
 CC amino acids 187-203 of human SNAP25) to a sample containing the
 CC botulinum toxin A so that hydrolysis of the peptide is initiated, then
 CC stopping hydrolysis of the peptide at different time points; and
 CC measuring the amount of hydrolysis at each time point by combining with a
 CC label capable of detecting free amino groups resulting from the
 CC hydrolysis. The amount of botulinum toxin A present in the sample is
 CC determined by comparing measurements with the amount of label produced
 CC from a known concentration of toxin measured under similar conditions.
 CC The method is useful for the quantitation of type A botulinum toxin.
 XX
 SQ Sequence 12 AA;
 Query Match 31.6%; Score 30; DB 20; Length 12;
 Best Local Similarity 54.5%; Pred. NO. 1.8e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 OY 1 DDADQRIIKQL 11
 I: I:||||
 DB 2 DEANQATKML 12
 RESULT 12
 AAU81241
 ID AAU81241 standard; Peptide; 15 AA.
 XX
 AC AAU81241;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Murine trkC antibody heavy chain CDR3 of variable region #2.
 XX
 KW Human; mouse; anti-trkC agonist monoclonal antibody; trkC; antibody;
 KW trkA; trkB; cisplatin-induced neuropathy; pyridoxine-induced neuropathy;
 KW peripheral neuropathy; diabetic neuropathy; neurodegenerative disease;
 KW large-fibre sensory neuropathy; amyotrophic lateral sclerosis; tumour;
 KW nerve cell injury; blood cell disorder; leukopaenia; eosinopaenia; wound;
 KW basopaenia; lymphopaenia; monocytopenia; neutropaenia; cancer; ulcer;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; diabetes;
 KW sickle cell disease; cardiac ischaemia; cerebrovascular disorder;
 KW cellular degeneration; gene therapy.
 XX
 OS Mus sp.
 XX
 PN WO200198361-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 22-JUN-2001; 2001WO-US20153.
 XX
 PR 22-JUN-2000; 2000US-213141P.
 PR 05-OCT-2000; 2000US-238319P.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Devaux B, Hongo JS, Presta LG, Shelton DL;
 XX
 DR WPI; 2002-130790/17.
 XX
 PT Novel anti-trkC agonist monoclonal antibody useful for treating
 PT neurodegenerative disease, shows no significant cross-reactivity with
 PT trkA/trkB, and recognizes epitope in domain 5 of trkC
 XX
 PS Claim 12; Page 58; 121pp; English.
 XX
 CC The invention relates to an anti-trkC agonist monoclonal antibody which
 CC shows no significant cross-reactivity with trkA or trkB, and recognizes

CC an epitope in domain 5 of trkC. The antibodies of the invention are
 CC effective in the treatment of cisplatin- or pyridoxine-induced
 CC neuropathy, peripheral neuropathy, diabetic neuropathy and large-fibre
 CC sensory neuropathy, neurodegenerative disease including amyotrophic
 CC lateral sclerosis, nerve cell injuries, disorders of insufficient blood
 CC cells such as leukaemia including eosinopaenia, basopaenia,
 CC lymphopaenia, monocytopenia, neutropenia, Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease and tumours. The sequences are
 CC also useful for inducing angiogenesis for treating wounds, ulcers and
 CC diabetic complications of sickle cell disease, for treating cardiac
 CC ischaemia and cerebrovascular disorders and in the diagnosis of diseases
 CC involving cellular degeneration. Sequences AA081229-AA081284 represent
 CC human and mouse anti-trkC agonist monoclonal antibodies and antibody
 CC fragments of the invention.

XX Sequence 15 AA;

Query Match 31.6%; Score 30; DB 23; Length 15;

Best Local Similarity 56.7%; Pred. No. 2.3e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFEGN 17

Db :||:|

1 KYIYGN 6

RESULT 13

AA44021

ID AAY44021 standard; peptide; 17 AA.

XX AC AAY44021;

XX DT 18-JAN-2000 (first entry)

XX DE Amino acids 187-203 of human SNAP25.

XX Enzymatic assay; quantitation; type A botulinum neurotoxin; proteolysis;
 KW fluorescamine; detection; human; synaptosomal protein; SNAP25;
 KW hydrolysis; amino group.

XX OS Homo sapiens.

XX PN US5965699-A.

XX PD 12-OCT-1999.

XX PF 06-NOV-1996; 96US-0743894.

XX PR 06-NOV-1996; 96US-0743894.

XX PA (USSA) US SEC OF ARMY.

XX PI Bostian KA, Schmidt JJ;

XX DR WPI; 1999-579939/49.

XX PT Quantitation of type A botulinum toxin -

XX PS Claim 1; Column 4; 28pp; English.

XX The invention relates to an enzymatic assay for the quantitation of
 CC type A botulinum toxin, by determining the proteolytic activity of
 CC botulinum neurotoxin type A using fluorescamine detection. The method
 CC comprises adding an analogue (e.g. AAY44022-Y44076) of this peptide
 CC (which represents amino acids 187-203 of the human synaptosomal protein
 CC SNAP25) to a sample containing the botulinum toxin A so that hydrolysis
 CC of the peptide is initiated, then stopping hydrolysis of the peptide at
 CC different time points; and measuring the amount of hydrolysis at each
 CC time point by combining with a label capable of detecting free amino
 CC groups resulting from the hydrolysis. The amount of botulinum toxin A
 CC present in the sample is determined by comparing measurements with the
 CC amount of label produced from a known concentration of toxin measured
 CC under similar conditions. The method is useful for the quantitation of

CC type A botulinum toxin.

XX Sequence 17 AA;

Query Match 31.6%; Score 30; DB 20; Length 17;

Best Local Similarity 54.5%; Pred. No. 2.7e+02;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDADQRIIKQL 11

Db :||:|

7 DEANQRATKML 17

RESULT 14

AA44039

ID AAY44039 standard; peptide; 17 AA.

XX AC AAY44039;

XX DT 18-JAN-2000 (first entry)

XX DE Human SNAP25 (amino acids 187-203) analogue #18.

XX Enzymatic assay; quantitation; type A botulinum neurotoxin; proteolysis;
 KW fluorescamine; detection; human; synaptosomal protein; SNAP25;
 KW hydrolysis; amino group.

XX OS Homo sapiens.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Modified-site 4 /label= Abu

XX PN US5965699-A.

XX PD 12-OCT-1999.

XX PF 06-NOV-1996; 96US-0743894.

XX PR 06-NOV-1996; 96US-0743894.

XX PA (USSA) US SEC OF ARMY.

XX PI Bostian KA, Schmidt JJ;

XX DR WPI; 1999-579939/49.

XX PT Quantitation of type A botulinum toxin -

XX PS Disclosure; Column 7-8; 28pp; English.

XX The invention relates to an enzymatic assay for the quantitation of
 CC type A botulinum toxin, by determining the proteolytic activity of
 CC botulinum neurotoxin type A using fluorescamine detection. Botulinum
 CC toxin A has been shown to cleave the synaptosomal neurotransmitter
 CC peptide SNAP25 between residues 197-198. The method comprises adding
 CC an analogue (e.g. AAY44022-Y44076) of the SNAP25 peptide (AAY44021,
 CC amino acids 187-203 of human SNAP25) to a sample containing the
 CC botulinum toxin A so that hydrolysis of the peptide is initiated, then
 CC stopping hydrolysis of the peptide at different time points; and
 CC measuring the amount of hydrolysis at each time point by combining with a
 CC label capable of detecting free amino groups resulting from the
 CC hydrolysis. The amount of botulinum toxin A present in the sample is
 CC determined by comparing measurements with the amount of label produced
 CC from a known concentration of toxin measured under similar conditions.
 CC The method is useful for the quantitation of type A botulinum toxin.

XX Sequence 17 AA;

Query Match 31.6%; Score 30; DB 20; Length 17;

Best Local Similarity 54.5%; Pred. No. 2.7e+02;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Job time : 28.2022 secs

QY 1 DDADQRIKQL 11
|:|:|:|:|
Db 7 DEANQRAKML 17

RESULT 15

AAV44045
ID AAY44045 standard; peptide; 17 AA.
XX AC AAY44045;
XX DT 18-JAN-2000 (first entry)
XX DE Human SNAP25 (amino acids 187-203) analogue T14B.
XX KW Enzymatic assay; quantitation; type A botulinum neurotoxin; proteolysis;
KW fluorescamine; detection; human; synaptosomal protein; SNAP25;
KW hydrolysis; amino group.
XX OS Homo sapiens.
OS Synthetic.

XX FH Key Location/Qualifiers
FT Modified-site 14
FT FT /label= Abu

XX PN US5965699-A.
XX PD 12-OCT-1999.
XX PF 06-NOV-1996; 96US-0743894.
XX PR 06-NOV-1996; 96US-0743894.

XX PA (USSA) US SEC OF ARMY.
XX PI Bostian KA, Schmidt JJ;
XX DR WPI; 1999-579939/49.

XX Quantitation of type A botulinum toxin -
XX Disclosure; Column 7-8; 28pp; English.

XX The invention relates to an enzymatic assay for the quantitation of
CC type A botulinum toxin, by determining the proteolytic activity of
CC botulinum neurotoxin type A using fluorescamine detection. Botulinum
CC toxin A has been shown to cleave the synaptosomal neurotransmitter
CC peptide SNAP25 between residues 197-198. The method comprises adding
CC an analogue (e.g. AAY4022-Y44076) of the SNAP25 peptide (AAY4021,
CC amino acids 187-203 of human SNAP25) to a sample containing the
CC botulinum toxin A so that hydrolysis of the peptide is initiated, then
CC stopping hydrolysis of the peptide at different time points; and
CC measuring the amount of hydrolysis at each time point by combining with a
CC label capable of detecting free amino groups resulting from the
CC hydrolysis. The amount of botulinum toxin A present in the sample is
CC determined by comparing measurements with the amount of label produced
CC from a known concentration of toxin measured under similar conditions.
CC The method is useful for the quantitation of type A botulinum toxin.

XX SQ Sequence 17 AA;

Query Match 31.6%; Score 30; DB 20; Length 17;
Best Local Similarity 54.5%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 DDADQRIKQL 11
|:|:|:|:|
Db 7 DEANQRAKML 17

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:21:27 ; Search time 29.4607 Seconds
(without alignments)
132.885 Million cell updates/sec

Title: US-09-836-073-16
Perfect score: 104
Sequence: 1 AALEAKICHQIEEYFGDF 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84.5	81.2	381	11 Q9CYB9	Q9cyb9 mus musculus
2	58.5	56.2	206	13 Q8QH15	Q8qhi5 gallus gall
3	55	52.9	135	16 Q87676	Q87676 aquifex ae
4	47	45.2	187	16 Q8Y2F2	Q8yzf2 anabaena sp
5	47	45.2	357	10 Q98SC2	Q98sc2 guillardia
6	47	45.2	448	3 Q42907	Q42907 schizosacch
7	46	44.2	284	3 Q04639	Q04639 saccharomyc
8	46	44.2	456	17 Q29131	Q29131 archaeoglob
9	45.5	43.8	569	11 Q8VDR9	Q8vdr9 mus musculus
10	45.5	43.8	844	11 Q9CUE3	Q9cue3 mus musculus
11	45.5	43.8	860	4 Q96HP0	Q96hp0 homo sapien
12	45	43.3	409	8 Q98RP4	Q98rp4 guillardia
13	45	43.3	928	10 Q9LJ02	Q9lj02 oryza sativ
14	45	43.3	1156	12 Q57230	Q57230 vaccinia vi
15	45	43.3	1164	12 Q90031	Q90031 variola vlr
16	45	43.3	1164	12 Q9JF79	Q9jf79 vaccinia vi

17	45	43.3	1164	12	Q90027	Q90027 variola maj
18	45	43.3	1164	12	Q8V4V3	Q8v4v3 monkeypox v
19	45	43.3	1164	12	Q8V2N1	Q8v2n1 camelppox vi
20	44	42.3	281	11	Q9CQK3	Q9cqk3 mus musculus
21	44	42.3	592	5	Q9VN01	Q9vn01 drosophila
22	43.5	41.8	119	10	Q9ZP89	Q9zpq5 neisseria m
23	43	41.3	88	16	Q9JQP5	Q9jqp5 polyandroca
24	43	41.3	185	5	Q9UAF3	Q9uaf3 polyandroca
25	43	41.3	297	4	Q9H1U9	Q9h1u9 homo sapien
26	43	41.3	427	16	Q05874	Q05874 mycobacteri
27	43	41.3	447	10	Q93V61	Q93v61 arabidopsis
28	43	41.3	510	10	Q43819	Q43819 pisum sativ
29	43	41.3	606	16	Q8REH6	Q8reh6 fusobacteri
30	43	41.3	647	5	Q9CT49	Q9ct49 trypanosoma
31	43	41.3	647	5	Q9GT48	Q9gt48 trypanosoma
32	43	41.3	647	5	Q9SWL5	Q9swl5 trypanosoma
33	43	41.3	723	16	Q8RH48	Q8rh48 fusobacteri
34	42.5	40.9	864	2	Q68395	Q68395 thauera aro
35	42.5	40.9	864	2	Q8VPT7	Q8vpt7 azoarcus sp
36	42	40.4	297	12	Q56780	Q56780 european ba
37	42	40.4	386	16	Q9CE77	Q9ce77 lactococcus
38	42	40.4	400	8	Q95BX4	Q95bx4 pteroceltis
39	42	40.4	436	17	Q9HK20	Q9hk20 thermoplas
40	42	40.4	468	5	Q9XV51	Q9xv51 caenorhabdi
41	42	40.4	541	16	Q9ZKY5	Q9zky5 helicobacte
42	42	40.4	542	16	Q25534	Q25534 helicobacte
43	42	40.4	905	3	Q13955	Q13955 schizosacch
44	42	40.4	969	16	Q9RZ15	Q9rzi5 deinococcus
45	42	40.4	1037	16	Q8XE42	Q8xeh2 escherichia

ALIGNMENTS

RESULT 1
Q9CYB9 PRELIMINARY; PRT; 381 AA.
AC Q9CYB9; (TREMREL. 17, Created)
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE Sjogren syndrome antigen B.
GN SSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE-EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK017822; BAB30957.1; -
MGD; MGI:98423; Ssb.

DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rrm; 1.
 DR PRINTS; PR00302; LUPUSLA.
 SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;

Query Match 81.2%; Score 84.5; DB 11; Length 381;
 Best Local Similarity 94.4%; Pred. No. 7.3e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2 ALEAKICHQIEYYFGDF 19
 ID Q8QHI5 PRELIMINARY; PRT; 206 AA.
 AC Q8QHI5;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Untranslated region binding-protein.
 GN USP.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 [1]
 SEQUENCE FROM N.A.
 RA L'Ecuver T.J., Fang H.-L.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF167897; AAL76289.1; -;
 SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFFB90E9 CRC64;

Query Match 56.2%; Score 58.5; DB 13; Length 206;
 Best Local Similarity 75.0%; Pred. No. 0.082;
 Matches 12; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 3 LEAKICHQIEYYFGD 18
 ID Q8QHI5 PRELIMINARY; PRT; 135 AA.
 AC Q8QHI5;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE THIOREDOXIN.
 GN TRX2 OR AQ_1811.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OC NCBI_TaxID=63363;
 [1]
 SEQUENCE FROM N.A.
 RA STRAIN-VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RL Nature 392:353-358(1998).
 DR EMBL; AE000757; AAC07635.1; -;
 DR InterPro; IPR000063; ThioRed.
 KW Complete proteome.
 SQ SEQUENCE 135 AA; 15746 MW; B9B8F51A91D7DD54 CRC64;

Query Match 52.9%; Score 55; DB 16; Length 135;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 8 CHQIEEYFGD 18
 ID Q8YZF2 PRELIMINARY; PRT; 187 AA.
 AC Q8YZF2;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical protein Alr0508.
 GN ALR0508.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OC NCBI_TaxID=103690;
 [1]
 SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriuchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL; AP003582; BAB72466.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 187 AA; 21553 MW; 77DB2FC9DD03A53E CRC64;

Query Match 45.2%; Score 47; DB 16; Length 187;
 Best Local Similarity 50.0%; Pred. No. 6.1;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIEEYFGD 18
 ID Q8YZF2 PRELIMINARY; PRT; 357 AA.
 AC Q8YZF2;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 42.4 kDa protein orf357 from chromosome 3.
 GN ORF357.
 OS Guillardia theta (Cryptomonas phi).
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
 OC NCBI_TaxID=55529;
 [1]
 SEQUENCE FROM N.A.
 RX MEDLINE=21233349; PubMed=11323671;
 RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,
 RA Wu X., Reith M., Cavalier-Smith T., Maier U.G.;
 RT "The highly reduced genome of an enslaved algal nucleus.";
 RL Nature 410:1091-1096(2001).
 DR EMBL; AF083031; AAK39661.1; -;
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 1.
 DR SMART; SM00320; WD40; 3.
 KW Hypothetical protein; Repeat; WD repeat.
 SQ SEQUENCE 357 AA; 42401 MW; 71AEB9F896A04C82 CRC64;

Query Match 45.2%; Score 47; DB 10; Length 357;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 97 ICHELESWFLGD 108
 ID Q8YZF2 PRELIMINARY; PRT; 357 AA.
 AC Q8YZF2;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 42.4 kDa protein orf357 from chromosome 3.
 GN ORF357.
 OS Guillardia theta (Cryptomonas phi).
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
 OC NCBI_TaxID=55529;
 [1]
 SEQUENCE FROM N.A.
 RX MEDLINE=21233349; PubMed=11323671;
 RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,
 RA Wu X., Reith M., Cavalier-Smith T., Maier U.G.;
 RT "The highly reduced genome of an enslaved algal nucleus.";
 RL Nature 410:1091-1096(2001).
 DR EMBL; AF083031; AAK39661.1; -;
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 1.
 DR SMART; SM00320; WD40; 3.
 KW Hypothetical protein; Repeat; WD repeat.
 SQ SEQUENCE 357 AA; 42401 MW; 71AEB9F896A04C82 CRC64;

Query Match 45.2%; Score 47; DB 10; Length 357;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 97 ICHELESWFLGD 108
 ID Q8YZF2 PRELIMINARY; PRT; 357 AA.
 AC Q8YZF2;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 42.4 kDa protein orf357 from chromosome 3.
 GN ORF357.
 OS Guillardia theta (Cryptomonas phi).
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
 OC NCBI_TaxID=55529;
 [1]
 SEQUENCE FROM N.A.
 RX MEDLINE=21233349; PubMed=11323671;
 RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,
 RA Wu X., Reith M., Cavalier-Smith T., Maier U.G.;
 RT "The highly reduced genome of an enslaved algal nucleus.";
 RL Nature 410:1091-1096(2001).
 DR EMBL; AF083031; AAK39661.1; -;
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 1.
 DR SMART; SM00320; WD40; 3.
 KW Hypothetical protein; Repeat; WD repeat.
 SQ SEQUENCE 357 AA; 42401 MW; 71AEB9F896A04C82 CRC64;

Query Match 45.2%; Score 47; DB 10; Length 357;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 97 ICHELESWFLGD 108
 ID Q8YZF2 PRELIMINARY; PRT; 357 AA.
 AC Q8YZF2;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 42.4 kDa protein orf357 from chromosome 3.
 GN ORF357.
 OS Guillardia theta (Cryptomonas phi).
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
 OC NCBI_TaxID=55529;
 [1]
 SEQUENCE FROM N.A.
 RX MEDLINE=21233349; PubMed=11323671;
 RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,
 RA Wu X., Reith M., Cavalier-Smith T., Maier U.G.;
 RT "The highly reduced genome of an enslaved algal nucleus.";
 RL Nature 410:1091-1096(2001).
 DR EMBL; AF083031; AAK39661.1; -;
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 1.
 DR SMART; SM00320; WD40; 3.
 KW Hypothetical protein; Repeat; WD repeat.
 SQ SEQUENCE 357 AA; 42401 MW; 71AEB9F896A04C82 CRC64;

Query Match 45.2%; Score 47; DB 10; Length 357;
 Best Local Similarity 50.0%; Pred. No. 12;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 97 ICHELESWFLGD 108
 ID Q8YZF2 PRELIMINARY; PRT; 357 AA.
 AC Q8YZF2;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical 42.4 kDa protein orf357 from chromosome 3.
 GN ORF357.
 OS Guillardia theta (Cryptomonas phi).
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.
 OC NCBI_TaxID=55529;
 [1]
 SEQUENCE FROM N.A.
 RX MEDLINE=21233349; PubMed=11323671;
 RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,
 RA Wu X., Reith M., Cavalier-Smith T., Maier U.G.;
 RT "The highly reduced genome of an enslaved algal nucleus.";
 RL Nature 410:1091-1096(2001).
 DR EMBL; AF083031; AAK39661.1; -;
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 1.
 DR SMART; SM00320; WD40; 3.
 KW Hypothetical protein; Repeat; WD repeat.
 SQ SEQUENCE 357 AA; 42401 MW; 71AEB9F896A04C82 CRC64;

QY 3 LEAKICHQIEEYFGD 18
|: |: |: |: |: |: |
DB 242 LKNKVSQMKNKYFGD 257

RESULT	6
042907	
ID	PRELIMINARY; PRT; 448 AA.
AC	O42907;
DC	01-JUN-1998 (TrEMBLrel. 06, Created)
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE	01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE	Hypothetical 51.2 kDa protein C119.16C in chromosome II.
GN	SPBC119.16C.
OS	Schizosaccharomyces pombe (Fission yeast).
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;
CC	Schizosaccharomycetes.
OX	NCBI_TaxID=4896;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=972;
RA	Wood V., Rajandream M.A., Barrell B.G., Skelton J., Chur
RL	Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC	-!- SIMILARITY: SOME, TO YEAST YKLO47W.
DR	EMBL; AL022117; CAAT1931.1; -.
CC	Hypothetical protein.
SK	SEQUENCE 448 AA: 51189 MW; 3E56729FA1A98DAB CRC64;

```

RESULT 7
Q04639
ID ID Q04639 PRELIMINARY; PRT; 284 AA.
AC Q04639;
DT 01-NOV-1996 (TremBrel. 01, Created)
DT 01-NOV-1996 (TremBrel. 01, Last sequence update)
DT 01-JUN-2002 (TremBrel. 21, Last annotation update)
DE Chromosome XIII COSMID 9745.
GN ERV41 OR YML067C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;.
RN [1]
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN-AB972;
RA Bowman S., Churcher C.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AB972;
RA Barrell B., Rajandream M.A.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z38114; CAA86253.1; -.
DR SGD; S0004532; ERV41.
SQ SEQUENCE 284 AA; 32882 MW; 200EEB06A92A3CCF CRC64;

```

RESULT 8	
O29131	PRELIMINARY; PRT; 456 AA.
ID	O29131
AC	O29131;
DT	01-JAN-1998 (TrEMBLrel. 05, Created)
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE	Hypothetical protein AF1134.
GN	AF1134.
OS	Archaeoglobus fulgidus.
OC	Archaea; Euryarchaeota; Archaeoglobales;
OC	Archaeoglobaceae; Archaeoglobus.
OX	NCBI_TaxID=2234;
RN	[1]
SEQUENCE FROM N.A.	
RC	STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX	MEDLINE=98049343; PubMed=9389475;
RA	Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA	Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,
RA	Richardson D.L., Kerlavage A.R., Graham D.E., Kyriades N.C.,
RA	Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA	Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA	Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA	Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uitterback T.,
RA	Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA	Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA	Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA	Venter J.C.;
RT	"The complete genome sequence of the hyperthermophilic, sulphate-
RT	reducing archaeon Archaeoglobus fulgidus.";
RL	Nature 390:364-370(1997).
RE	EMBL; AE001026; AAB90123.1; -.
DR	TIGR; AF1134; -.
DR	InterPro; IPR000379; Ser_estrs_site.
DR	Hypothetical protein; Complete proteome.
SW	SEQUENCE 456 AA; 51748 MW; 37CB62B58C2C9357 CRC64;

```

RESULT 9
Q8VDR9 Q8VDR9 PRELIMINARY; PRT; 569 AA.
AC Q8VDR9;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Hypothetical 65.3 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021414; AAH21414.1; -.
KW Hypothetical protein.
FT NON_TER .1
SQ SEQUENCE 569 AA; 65306 MW; 40AC4726813974B1 CRC64;

Query Match 43.8%; Score 45.5; DB 11; Length 569;
Best Local Similarity 52.9%; Pred. No. 35;
Matches 4; Mismatches 1; Indels 3; Gaps 1;

```

[illegible]

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:42 ; Search time 5.23034 Seconds

(without alignments)
150.669 Million cell updates/sec

Title: US-09-836-073-16

Perfect score: 104

Sequence: 1 AALEAKICHQIEEYFGDF 19

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88.5	85.1	404	1 LA_BOVIN	P10881 bos taurus
2	88.5	85.1	408	1 LA_HUMAN	P05455 homo sapien
3	88.5	85.1	415	1 LA_RAT	P38556 rattus norv
4	84.5	81.2	415	1 LA_MOUSE	P32067 mus musculus
5	59.5	57.2	427	1 LAB_XENLA	P28049 xenopus lae
6	55.5	53.4	428	1 LAB_XENLA	P28048 xenopus lae
7	47	45.2	482	1 LBP_RABIT	P17454 oryctolagus
8	46	44.2	352	1 YMG7_YEAST	Q04651 saccharomyc
9	45	43.3	620	1 ACE1_CAEER	Q27459 caenorhabdi
10	45	43.3	620	1 ACE1_CAEEL	P38433 caenorhabdi
11	45	43.3	1164	1 RPO2_COWPX	P17474 cowpox viru
12	45	43.3	1164	1 RPO2_VACCV	P19798 vaccinia vi
13	45	43.3	1164	1 RPO2_VARV	P33811 variola vir
14	44.5	42.8	383	1 LA_AEDAL	Q26457 aedes albop
15	44	42.3	552	1 ESTE_MYZPE	P35501 myzus persi
16	44	42.3	564	1 ESTE_MYZPE	P35502 myzus persi
17	43	41.3	609	1 YA38_METJA	Q58458 methanococc
18	42	40.4	421	1 HMDH_AERPE	Q9yasa4 aeropyrum p
19	42	40.4	428	1 Y813_TREPA	O83789 treponema p
20	42	40.4	1037	1 ACRD_ECOLI	P24177 escherichia
21	41	39.4	263	1 MTX2_MOUSE	O88441 mus musculu
22	41	39.4	926	1 MEI9_DROME	Q24087 drosophila
23	40.5	38.9	390	1 LA_DROME	P40796 drosophila
24	40	38.5	263	1 MTX2_HUMAN	O75431 homo sapien
25	40	38.5	320	1 RLAO_SOYBN	P50346 glycine max
26	40	38.5	509	1 STK_HYDAT	P17713 hydra atten
27	40	38.5	571	1 PTI_CHLPN	Q92983 chlamydia p
28	40	38.5	597	1 MBHL_RHOCA	P15284 rhodobacter
29	40	38.5	812	1 MCM3_MOUSE	P25206 mus musculu
30	40	38.5	919	1 RPO2_CAPVK	P16716 capripoxvir
31	40	38.5	1253	1 ROLF_SFV	P03315 senliki for
32	40	38.5	1505	1 A77B_SHEEP	Q9xt50 ovvis aries
33	39	37.5	282	1 DAAA_STAHA	P54694 staphylococ

34 39 37.5 352 1 HAO3_RAT Q07523 rattus norv
35 39 37.5 396 1 PRRC_ECOLI P17223 escherichia
36 39 37.5 409 1 HEM1_THEVO Q97868 thermoplasm
37 39 37.5 477 1 ARP7_YEAST Q12406 saccharomyc
38 39 37.5 503 1 TCPT_VIBCH P29480 vibrio chol
39 39 37.5 512 1 ACH3_CARAU P18845 carassius a
40 39 37.5 599 1 HM21_HUMAN Q99728 homo sapien
41 39 37.5 777 1 BAR1_HUMAN Q99728 homo sapien
42 39 37.5 808 1 MCM3_HUMAN P25206 mus musculu
43 39 37.5 827 1 YNH4_CAEEL P32742 caenorhabdi
44 39 37.5 1147 1 MYSB_ACACA P19706 acanthamoeb
45 39 37.5 1727 1 ALM1_SCHPO Q9utk5 schizosacch

ALIGNMENTS

RESULT 1
LA_BOVIN STANDARD; PRT; 404 AA.
AC P10881;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
DE homolog).
GN SSB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=89202037; PubMed=2468131;
RA Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
sequences for RNA-binding";
RL Nucleic Acids Res. 17:2233-2244(1989).
CC -!- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
CC AND 7-2 RNAs.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE
CC C-TERMINAL PART OF THE PROTEIN
CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13698; CAA31986.1; -
DR F03849; S03849.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
KW RNA-binding; Nuclear protein.
FT DOMAIN 111 187 RNA-BINDING (RRM).
SQ SEQUENCE 404 AA; 46534 MW; 4EE30B5C262AD6A1 CRC64;

Query Match

85.1%; Score 88.5; DB 1; Length 404;

Best Local Similarity 94.7%; Pred. No. 1.2e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEEYFGDF 19
DB 11 AALEAKICHQI-EYFGDF 28

RESULT 2
LA_HUMAN STANDARD; PRT; 408 AA.
AC P05455;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lupus La protein (Sjogren syndrome type B antigen (SS-B)) (La
DE ribonucleoprotein) (La autoantigen).
GN SSB.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89202037; PubMed=2458131;
RX Chan E.K.L., Sullivan K.F., Tan E.M.;
RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus
RT sequences for RNA-binding.";
RL Nucleic Acids Res. 17:2233-2244(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89053970; PubMed=3192525;
RX Chambers J.C., Kenan D., Martin B.J., Keene J.D.;
RA "Genomic structure and amino acid sequence domains of the human La
RT autoantigen.";
RL J. Biol. Chem. 263:18043-18051(1988).
RN [3]
RP SEQUENCE OF 54-408 FROM N.A.
RX MEDLINE=88199081; PubMed=2452201;
RX Sturgess A.D., Peterson M.G., McNeillage L.J., Whittingham S.,
RA Coppel R.S.;
RT "Characteristics and epitope mapping of a cloned human autoantigen
RT La.";
RL J. Immunol. 140:3212-3218(1988).
RN [4]
RP SEQUENCE OF 54-97 FROM N.A.
RX MEDLINE=85166283; PubMed=3856888;
RX Chambers J.C., Keene J.D.;
RA "Isolation and analysis of cDNA clones expressing human lupus La
RT antigen.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2115-2119(1985).
RN [5]
RP FUNCTION.
RX MEDLINE=89251617; PubMed=2470590;
RX Gottlieb E., Steitz J.A.;
RA "Function of the mammalian La protein: evidence for its action in
RT transcription termination by RNA polymerase III.";
RL EMBO J. 8:851-861(1989).
RN [6]
RP PHOSPHORYLATION.
RX MEDLINE=97207017; PubMed=9054510;
RX Fan H., Sakulich A.L., Goodier J.L., Zhang X., Qin J., Marale R.J.;
RA "Phosphorylation of the human La antigen on serine 366 can regulate
RT recycling of RNA polymerase III transcription complexes.";
RL Cell 88:707-715(1997).
RN [7]
RP FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC OF RNA POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC AND 7-2 RNAs.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE

CC C-TERMINAL PART OF THE PROTEIN.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.
CC -1- DISEASE: SERA FROM PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS
CC OFTEN CONTAIN ANTIBODIES THAT REACT WITH THE NORMAL CELLULAR
CC LA PROTEIN AS IF THESE ANTIGEN WAS FOREIGN.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X13697; CAA31985.1; -
CC EMBL; J04205; AAA51885.1; -
CC PIR; A31888; A31888.
CC PIR; A22956; A22956.
CC PIR; A31273; A31273.
CC PIR; S03848; S03848.
CC PIR; S11013; S11013.
CC Genew; HGNC:11316; SSB.
CC MTM; 109090;
CC InterPro; IPR002344; Lupus_La.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS0102; RRM; 1.
CC PROSITE; PS00030; RRM_LNP_1; 1.
CC Systemic lupus erythematosus; RNA-binding; Phosphorylation;
CC Nuclear protein.
CC DOMA1N 111 187 RNA-BINDING (RRM).
CC MOD_RES 366 366 PHOSPHORYLATION (BY CK2).
CC SEQUENCE 408 AA; 46837 MW; EC153C15F9187FC4 CRC64;
CC
CC Query Match 85.1%; Score 88.5; DB 1; Length 408;
CC Best Local Similarity 94.7%; Pred. No. 1.2e-06;
CC Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
CC
CC QY 1 AALEAKICHQIEEYFGDF 19
CC DB 11 AALEAKICHQI-EYFGDF 28
CC
CC RESULT 3
CC LA_RAT STANDARD; PRT; 415 AA.
CC ID LA_RAT
CC AC P38656;
CC DT 01-FEB-1995 (Rel. 31, Created)
CC DT 01-FEB-1995 (Rel. 31, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen
CC homolog).
CC DE SSB OR SS-B.
CC OS Rattus norvegicus (Rat).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
CC OX NCBI_TaxID=10116;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=93246255; PubMed=7916708;
CC RA Semsel I., Troester H., Bartsch H., Schwemmler M., Igloi G.L.,
CC Bachmann M.;
CC "Isolation of rat cDNA clones coding for the autoantigen SS-B/La:
CC detection of species-specific variations.";
CC RL Gene 126:265-268(1993).
CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
CC AND 7-2 RNAs.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE

CC AND 7-2 RNAs.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X67859; CAA48043.1; -
 CC PIR; JCI1494;
 CC InterPro: IPR002344; Lupus.La.
 CC InterPro: IPR000504; RNA_rec_mot.
 CC Pfam: PF00076; rtm: 1;
 CC PRINTS; PR00302; LUPUSLA.
 CC SMART; SM00360; RRM; 1.
 CC PROSITE; PS0102; RRM; 1.
 CC PROSITE; PS00030; RRM_RNP_1; 1.
 CC RNA-binding; Nuclear protein; Phosphorylation.
 KW RNA-binding; Nuclear protein; RNA-BINDING (RRM).
 FT DOMAIN 111 187
 FT SEQUENCE 415 AA; 47777 MW; 033FD9CC1B475F98 CRC64;
 CC
 CC Query Match 85.1%; Score 88.5; DB 1; Length 415;
 CC Best Local Similarity 94.7%; Pred. No. 1.3e-06;
 CC Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 CC
 CC QY 1 ALEAKICHQIEYYFGDF 19
 CC LA_MOUSE
 CC ID LAB_MOUSE STANDARD; PRT; 415 AA.
 CC AC P32067;
 CC DT 01-OCT-1993 (Rel. 27, Created)
 CC DT 01-OCT-1993 (Rel. 27, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Lupus La protein homolog (La ribonucleoprotein). (La autoantigen
 CC DE homolog).
 CC GN SSB OR SS-B.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC OX NCBI_TaxID=10090;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE-93203630; PubMed-8454877;
 CC RA Topfer F., Gordon T., McCluskey J.;
 CC RT "Characterization of the mouse autoantigen La (SS-B). Identification
 CC of conserved RNA-binding motifs, a putative ATP binding site and
 CC reactivity of recombinant protein with poly(U) and human
 CC autoantibodies.";
 CC RT J. Immunol. 150:3091-3100(1993).
 CC RL [2]
 CC RN SEQUENCE OF 1-11 FROM N.A.
 CC RP Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC RL Grodz D., Bachmann M.;
 CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
 CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
 CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
 CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
 CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
 CC AND 7-2 RNAs.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L00993; AAA39415.1; -
 CC EMBL; Y07951; CAA69249.1; -
 CC MGD; MGI:98423; Ssb.
 CC InterPro: IPR002344; Lupus.La.
 CC InterPro: IPR000504; RNA_rec_mot.
 CC Pfam: PF00076; rtm: 1;
 CC PRINTS; PR00302; LUPUSLA.
 CC SMART; SM00360; RRM; 1.
 CC PROSITE; PS0102; RRM; 1.
 CC PROSITE; PS00030; RRM_RNP_1; 1.
 CC RNA-binding; Nuclear protein; Phosphorylation.
 KW RNA-binding; Nuclear protein; RNA-BINDING (RRM).
 FT DOMAIN 111 187
 FT SEQUENCE 415 AA; 47756 MW; 2D75197692FDC933 CRC64;
 CC
 CC Query Match 81.2%; Score 84.5; DB 1; Length 415;
 CC Best Local Similarity 94.4%; Pred. No. 5.4e-06;
 CC Matches 17; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 CC
 CC QY 2 ALEAKICHQIEYYFGDF 19
 CC LAB_XENLA
 CC ID LAB_XENLA STANDARD; PRT; 427 AA.
 CC AC P28049;
 CC DT 01-AUG-1992 (Rel. 23, Created)
 CC DT 01-AUG-1992 (Rel. 23, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen
 CC DE homolog B).
 CC GN LAB1.
 CC OS Xenopus laevis (African clawed frog).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 CC OC Xenopodidae; Xenopus.
 CC OX NCBI_TaxID=8355;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Oocyte;
 CC RX MEDLINE-93287095; PubMed-8510143;
 CC RA Scherly D., Stutz F., Lin-Marq N., Clarkson S.G.;
 CC RT "La proteins from Xenopus laevis. cDNA cloning and developmental
 CC expression";
 CC RL J. Mol. Biol. 231:196-204(1993).
 CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
 CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
 CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
 CC POLYMERASE III TRANSCRIPTS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: BARELY DETECTABLE IN STAGE I/II OOCYTES,
 CC ACCUMULATE IN STAGE III/IV OOCYTES, THEN EXHIBIT A ROUGHLY
 CC CONSTANT STEADY STATE LEVEL IN MATURE OOCYTES, EGGS, AND EARLY
 CC EMBRYOS.
 CC -1- PTM: PHOSPHORYLATED (PROBABLE).
 CC -1- MISCELLANEOUS: THERE ARE TWO FORMS OF LA, LAA AND LAB, IN XENOPUS.
 CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

or send an email to license@lsb-sib.ch).

```

DR PIR; S33818; S33818.
DR InterPro; IPR002344; Lupus_La.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF000076; rrm; 1.
DR PRINTS; PR00302; LUPUSLA.
DR SMART; SM00360; RRM; 1..
DR PROSITE; PS00102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
DR KFT DOMAIN 111 203 RNA-BINDING; Nuclear protein; Phosphorylation.
DR KFT DOMAIN 196 212 RNA-BINDING (RRM).
DR FTFT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
DR FTFT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
DR SQ SEQUENCE 428 AA; 48864 MW; AEB3A38B7D2E3EC3 CRC64;

Query Match 53.4%; Score 55.5; DB 1; Length 428;
Best Local Similarity 73.3%; Pred. No. 0.21;
Matches 11; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 4 EAKICHQIEYYFGD 18
: ||| || |||||
DB 14 DTKICEOI-EYFQD 27

RESULT 7
LBP_RABIT STANDARD; PRT; 482 AA.
AC PI7454;
ID LBP_RABIT
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Lipopolysaccharide-binding protein precursor (LBP).
GN LBP.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90385281; PubMed=2402637;
RA Schumann R.R., Leong S.R., Flagg G.W., Gray P.W., Wright S.D.,
RA Mathison J.C., Tobias P.S., Ulevitch R.J.;
RT "Structure and function of lipopolysaccharide binding protein.";
RL Science 249:1429-1431(1990).
RN [2]
RP SEQUENCE OF 27-66.
RC TISSUE=Serum;
RX MEDLINE=86306528; PubMed=2427635;
RA Tobias P.S., Soldau K., Ulevitch R.J.;
RT "Isolation of a lipopolysaccharide-binding acute phase reactant from
RT rabbit serum."
RL J. Exp. Med. 164:777-793(1986).
CC -1- FUNCTION: BINDS TO THE LIPID A MOIETY OF BACTERIAL
CC LIPOPOLYSACCHARIDES (LPS), A GLYCOLIPID PRESENT IN THE OUTER
CC MEMBRANE OF ALL GRAM-NEGATIVE BACTERIA. THE LBP/LPS COMPLEX SEEMS
CC TO INTERACT WITH THE CD14 RECEPTOR.
CC -1- SIMILARITY: BELONGS TO THE BPI/CETP/LBP/PLTP FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M35534; AAA99235.1; -.
CC PIR; B35843; B35843.
CC HSSP; PI7213; 1BP1.
DR InterPro; IPR001124; LBP_BPI_CETP.
DR Pfam; PF01273; LBP_BPI_CETP; 1.
DR Pfam; PF02886; LBP_BPI_CETP-C; 1.
DR SMART; SM00328; BPI1; 1.
DR SMART; SM00329; BPI2; 1.

```


DR PROSITE; PS00400; LBP_BPI_CETP; 1.
KW Lipid transport; Antibiotic; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 25
FT CHAIN 27 482 LIPOPOLYSACCHARIDE-BINDING PROTEIN.
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 57 57 E -> G (IN REF. 2).
FT CONFLICT 63 63 S -> F (IN REF. 2).
SQ SEQUENCE 482 AA; 54001 MW; 628A6E0A647200C2 CRC64;

Query Match 45.2%; Score 47; DB 1; Length 482;
Best Local Similarity 81.8%; Pred. No. 5.2;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 LEAKICHOIEE 13
|:|:|:|:|:|
Db 194 LESKICHOIEE 204

RESULT 8
ID YMG7_YEAST STANDARD; PRT; 352 AA.
AC Q04651;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 40.7 kDa protein in DAK1-ORC1 intergenic region.
GN YML067C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Bowman S., Churcher C., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO YEAST YAL042W AND S.POMBE SPAC24B11.08C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z38114; CAA86254.1; -.
DR SGD; S0004532; YML067C.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 24 44 POTENTIAL.
FT TRANSMEM 290 310 POTENTIAL.
FT TRANSMEM 312 332 POTENTIAL.
SQ SEQUENCE 352 AA; 40705 MW; A9F002FB97666501 CRC64;

Query Match 44.2%; Score 46; DB 1; Length 352;
Best Local Similarity 56.2%; Pred. No. 5.1;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 4 EAKICHOIEEYFGDF 19
|:|:|:|:|:|
Db 196 ELKFNHVINFEFGDF 211

RESULT 9
ID ACBL_CABR STANDARD; PRT; 620 AA.
AC Q27459;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Acetylcholinesterase 1 precursor (EC 3.1.1.7) (AChE 1).
GN ACE-1.
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97069944; PubMed=8912924;
RA Grauso M., Culetto E., Berge J.B., Tountant J.P., Arpagaus M.;
RT 'Sequence comparison of ACE-1, the gene encoding acetylcholinesterase
RT of class A, in the two nematodes Caenorhabditis elegans and
RT Caenorhabditis briggsae.';
RL DNA Seq. 6:217-227(1996).
CC -1- FUNCTION: RAPIDLY HYDROLYZES CHOLINE RELEASED INTO THE SYNAPSE.
CC -1- CATALYTIC ACTIVITY: Acetylcholine + H(2)O -> choline + acetate.
CC -1- SUBUNIT: OLIGOMER COMPOSED OF DISULFIDE-LINKED HOMODIMERS (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MAY BE SECRETED OR MEMBRANE ASSOCIATED VIA
CC A NON-CATALYTIC SUBUNIT (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U41846; BAB41269.1; -.
DR HSP; P21836; LMAA.
DR InterPro; IPR002018; CarbesteraseB.
DR InterPro; IPR000997; Cholinesterase.
DR InterPro; IPR000379; Ser_estrs_site.
DR Pfam; PF001135; Coesterase_1.
DR PRINTS; PR00878; CHOLNESTRASE.
DR PROSITE; PS00122; CARBOXYLESTERASE_B_1; 1.
DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW Hydrolyase; Serine esterase; Synapse; Membrane; Signal; Glycoprotein;
KW Neurotransmitter degradation; Multigene family.
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 620 ACETYLCHOLINESTERASE 1.
FT ACT_SITE 216 216 BY SIMILARITY.
FT ACT_SITE 346 346 BY SIMILARITY.
FT ACT_SITE 468 468 BY SIMILARITY.
FT DISULFID 82 109 BY SIMILARITY.
FT DISULFID 270 286 BY SIMILARITY.
FT DISULFID 430 558 BY SIMILARITY.
FT DISULFID 618 618 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 486 486 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 620 AA; 71501 MW; 69D73CD3996E11FC CRC64;

Query Match 43.3%; Score 45; DB 1; Length 620;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 CHOTEYVFGDF 19
|:|:|:|:|:|
Db 82 CIQSDTYFGDF 93

RESULT 10
ID ACBL_CAEEL STANDARD; PRT; 620 AA.
AC P38433;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Acetylcholinesterase 1 precursor (EC 3.1.1.7) (AChE 1).


```
AC P19798;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase 132 kDa polypeptide (EC 2.7.7.6).
GN RPO132 OR A24R.
OS Vaccinia virus (strain WR), and
OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10254, 10249;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WR;
RX MEDLINE=91082452; Pubmed=1824607;
RA Megadzie B.Y., Holmes M.H., Cole N.B., Jones E.V., Earl P.L.,
RA Moss B.;
RT "Identification, sequence, and expression of the gene encoding the
RT second-largest subunit of the vaccinia virus DNA-dependent RNA
RT polymerase.";
RL Virology 180:88-98(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Copenhagen;
RX MEDLINE=91021027; Pubmed=2219722;
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RT "The complete DNA sequence of vaccinia virus.";
RL Virology 179:247-266(1990).
RN [3]
RP COMPLETE GENOME.
RC STRAIN=Copenhagen;
RX MEDLINE=91021027; Pubmed=2219722;
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RT "The complete DNA sequence of vaccinia virus'.";
RL Virology 179:517-563(1990).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THIS ENZYME CONSISTS OF AT LEAST EIGHT SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37415; AAA72882.1; -
DR EMBL; M35027; AAA48148.1; -
DR PIR; H42519; RNV28T.
DR PIR; A38517; RNVZMA.
DR InterPro; IPR001572; RNA_pol_B.
DR Pfam; PF00562; RNA_pol_B; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
DR Transferase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger.
FT ZN_FING 1087 1106 C4-TYPE (POTENTIAL).
SQ SEQUENCE 1164 AA; 133363 MW; B1F5C1484BA37D0D CRC64;
Query Match 43.3%; Score 45; DB 1; Length 1164;
Best Local Similarity 56.2%; Pred. No. 26;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 3 LEAKICHOIEEYFGD 18
DB 505 LEKKICEYIRSYKDD 520
RESULT 14
LA_AEDAL
ID LA_AEDAL STANDARD; PRT; 383 AA.
AC Q26457;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
OS Aedes albopictus (Forest day mosquito).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Culicoidae; Aedes.
OX NCBI_TaxID=7160;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96135233; Pubmed=8551578;
RA Pardigon N., Strauss J.H.;
```

```
RESULT 13
RPO2_VARV
ID RPO2_VARV STANDARD; PRT; 1164 AA.
AC P3811;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE DNA-directed RNA polymerase 132 kDa polypeptide (EC 2.7.7.6).
GN RPO132 OR A24R.
OS Variola virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10255;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=India-1967 / Isolate Ind3;
RX MEDLINE=93202281; Pubmed=8384129;
RA Shchelkunov S.N., Blinov V.M., Sandakhchiev L.S.;
RT "Genes of variola and vaccinia viruses necessary to overcome the host
RT protective mechanisms.";
RL FEBS Lett. 319:80-83(1993).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THIS ENZYME CONSISTS OF AT LEAST EIGHT SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69198; CAA49069.1; -
DR PIR; G36850; G36850.
DR InterPro; IPR001572; RNA_pol_B.
DR Pfam; PF00562; RNA_pol_B; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
DR Transferase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger.
FT ZN_FING 1087 1106 C4-TYPE (POTENTIAL).
SQ SEQUENCE 1164 AA; 133329 MW; 0452B84ED810CD53 CRC64;
Query Match 43.3%; Score 45; DB 1; Length 1164;
Best Local Similarity 56.2%; Pred. No. 26;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 3 LEAKICHOIEEYFGD 18
DB 505 LEKKICEYIRSYKDD 520
RESULT 14
LA_AEDAL
ID LA_AEDAL STANDARD; PRT; 383 AA.
AC Q26457;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
OS Aedes albopictus (Forest day mosquito).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Culicoidae; Aedes.
OX NCBI_TaxID=7160;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96135233; Pubmed=8551578;
RA Pardigon N., Strauss J.H.;
```

RT "Mosquito homolog of the La autoantigen binds to Sindbis virus RNA.";
 RL J. Virol. 70:1173-1181(1996).
 CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTION TERMINATION BY RNA
 CC POLYMERASE III. BINDS RNA AND DNA. BINDS TO THE 3' END OF THE
 CC MINUS STRAND OF SINDBIS VIRUS RNA. THIS MAY BE SIGNIFICANT FOR
 CC SINDBIS VIRUS RNA REPLICATION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. PRIMARILY NUCLEAR,, BUT SIGNIFICANT
 CC AMOUNTS ARE PRESENT IN THE CYTOPLASM.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: S80954; AAB35931.1;
 CC InterPro: IPR002344; Lupus_La.
 CC InterPro: IPR000504; RNA_rec_mot.
 CC Pfam: PF000076; rrm; 1.
 CC PRINTS: PR00302; LUPUSLA.
 CC SMART: SM00360; RRM; 1.
 CC PROSITE: PS0102; RRM; 1.
 CC PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
 CC RNA-binding; Nuclear protein; DNA-binding.
 CC DOMAIN 141 228 RNA-BINDING (RRM).
 CC SEQUENCE 383 AA; 44430 MW; 4E5CC8F21C40F452 CRC64;
 CC
 CC Query Match 42.8%; Score 44.5; DB 1; Length 383;
 CC Best Local Similarity 62.5%; Pred. No. 10;
 CC Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;
 CC
 CC YQ 3 LEAKICHOIEEYFGD 18
 CC III I:|||||
 CC Db 43 LEASTIRQL-EYIFGD 57
 CC
 CC RESULT 15
 CC ESTE.MVZPE STANDARD; PRT; 552 AA.
 CC ID P35501;
 CC DT 01-JUN-1994 (Rel. 29, Created)
 CC DT 01-JUN-1994 (Rel. 29, Last sequence update)
 CC DT 01-OCT-1994 (Rel. 30, Last annotation update)
 CC DE Esterase E4 precursor (EC 3.1.1.1) (Carboxylic-ester hydrolase).
 CC OS Myzus persicae (peach-potato aphid).
 CC OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha;
 CC Aphidiformes; Aphidoidea; Aphididae; Macrosiphini; Myzus.
 CC OX NCBI_TaxID=131164;
 CC RN [1]
 CC SEQUENCE FROM N.A., AND SEQUENCE OF 24-63.
 CC STRAIN-R3 / Isolate 794J;
 CC RX MEDLINE=93384534; PubMed=8373371;
 CC RA Field L.M., Williamson M.S., Moores G.D., Devonshire A.L.;
 CC RT "Cloning and analysis of the esterase genes conferring insecticide
 CC resistance in the peach-potato aphid, Myzus persicae (Sulzer).";
 CC RL Biochem. J. 294:569-574(1993)
 CC -1- FUNCTION: OVERPRODUCTION OF NONSPECIFIC ESTERASES IS A COMMON
 CC MECHANISM OF RESISTANCE TO ORGANOPHOSPHATE INSECTICIDES.
 CC -1- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O = an alcohol + a
 CC carboxylic anion.
 CC -1- MISCELLANEOUS: THIS ESTERASE CONFERS INSECTICIDE RESISTANCE.
 CC -1- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X74554; CAA52648.1;
 CC PIR: S36786; S36786.
 CC DR HSP; P21836; JMAA.
 CC DR InterPro: IPR002018; CarbesteraseB.
 CC DR InterPro: IPR000379; Ser_estra_site.
 CC Pfam: PF00135; Coesterase; 1.
 CC DR PROSITE: PS00122; CARBOXYLESTERASE_B_1; 1.
 CC DR PROSITE: PS00941; CARBOXYLESTERASE_B_2; FALSE_NEG.
 CC Hydrolase; Serine esterase; Glycoprotein; Signal.
 CC KW SIGNAL 1 23
 CC FT CHAIN 24 552
 CC FT ACT_SITE 214 214
 CC FT ACT_SITE 339 339
 CC FT ACT_SITE 463 463
 CC FT DISULFID 89 106
 CC FT DISULFID 266 277
 CC FT CARBOHYD 81 81
 CC FT CARBOHYD 269 269
 CC FT CARBOHYD 371 371
 CC FT CARBOHYD 404 404
 CC FT CARBOHYD 443 443
 CC SQ SEQUENCE 552 AA; 61348 MW; B97B67272DFF7209 CRC64;
 CC
 CC Query Match 42.3%; Score 44; DB 1; Length 552;
 CC Best Local Similarity 50.0%; Pred. No. 18;
 CC Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 CC
 CC YQ 3 LEAKICHOIEEYFGD 18
 CC I:|||||
 CC Db 379 LREKTAQDIKEFYFGD 394
 CC
 CC Search completed: April 23, 2003, 13:28:10
 CC Job time : 7.23034 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:03 ; Search time 10.6742 Seconds
(without alignments)
171.119 Million cell updates/sec

Title: US-09-836-073-16

Perfect score: 104
Sequence: 1 AALEAKICHQIEYYFGDF 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR73:*

1: Pirl:*

2: Pirl2:*

3: Pirl3:*

4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88.5	85.1	404	1 S03849	ribonucleoprotein
2	88.5	85.1	408	1 A31888	ribonucleoprotein
3	88.5	85.1	415	1 JCI1494	ribonucleoprotein
4	59.5	57.2	427	1 S33817	ribonucleoprotein
5	55.5	53.4	428	1 S33818	ribonucleoprotein
6	55	52.9	135	2 B70456	thioredoxin - Aquil
7	47	45.2	187	2 AC1870	hypothetical prote
8	47	45.2	357	2 H90120	hypothetical prote
9	47	45.2	448	2 T39314	hypothetical prote
10	47	45.2	482	2 B35843	lipopolysaccharide
11	46	44.2	352	2 S48331	probable membrane
12	46	44.2	456	2 E69391	hypothetical prote
13	45	43.3	409	2 H90096	eukaryotic release
14	45	43.3	620	2 A54413	acetylcholinestera
15	45	43.3	1156	2 T37411	RNA polymerase sub
16	45	43.3	1164	1 RN728T	DNA-directed RNA p
17	45	43.3	1164	1 RN728T	DNA-directed RNA p
18	45	43.3	1164	2 T28566	DNA-directed RNA p
19	45	43.3	1164	2 F72166	A25R protein - var
20	45	43.3	1164	2 G36850	A24R protein - var
21	44	42.3	552	2 S36786	carboxylesterase (
22	44	42.3	564	2 S36787	conserved hypothet
23	43	41.3	88	2 H81014	probable desA3 pro
24	43	41.3	427	2 G70590	glucose-1-phosphat
25	43	41.3	510	2 T06495	modulation factor
26	43	41.3	609	2 A64432	hypothetical prote
27	42	40.4	386	2 H86870	probable 3-hydroxy
28	42	40.4	421	2 E72573	hypothetical prote
29	42	40.4	428	2 B71278	hypothetical prote

30	42	40.4	468	2 T21535	hypothetical prote
31	42	40.4	541	2 H71887	hypothetical prote
32	42	40.4	542	2 G64627	hypothetical prote
33	42	40.4	905	2 T38314	probable vacuolar
34	42	40.4	969	2 A75634	McrB-related prote
35	42	40.4	1037	2 E65022	acriflavin resista
36	42	40.4	1037	2 D91045	aminoglycoside eff
37	42	40.4	1037	2 G85889	hypothetical prote
38	42	40.4	1069	2 AF1930	hypothetical prote
39	41.5	39.9	463	2 T39004	probable histone a
40	41	39.4	153	2 T19054	hypothetical prote
41	41	39.4	203	2 T28352	ORF MSV191 MTG mot
42	41	39.4	528	2 T02972	SRP1 protein homol
43	41	39.4	650	2 C69678	involved in polyke
44	41	39.4	771	2 T43612	transposase - vers
45	41	39.4	926	2 S58936	meiotic recombinat

ALIGNMENTS

RESULT 1

S03849

ribonucleoprotein La - bovine

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B

C:Species: Bos primigenius taurus (cattle)

C>Date: 31-Dec-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: S03849

R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.

Nucleic Acids Res. 17, 2233-2244, 1989

A:Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequ

A:Reference number: S03848; MUID:89202037; PMID:2468131

A:Accession: S03849

A:Molecule type: mRNA

A:Residues: 1-404 <CHA>

A:Cross-references: EMBL:X13698; NID:q755; PIDN:CAA31986.1; PID:g756

A:Note: part of this sequence was confirmed by protein sequencing

C:Comment: This protein associates with a variety of small RNA molecules, most of

ay act as a transcription termination factor.

C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

F:112-178/Domain: blocked amino end; phosphoprotein; RNA binding

F:113-118/Region: ribonucleoprotein repeat homology <RRM>

F:151-158/Region: RNA-binding RNP2 motif

F:228-404/Domain: phosphorylated #status predicted <PHY>

Query Match 85.1% Score 88.5; DB 1; Length 404;
Best Local Similarity 94.7%; Pred. No. 3.3e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
DB 11 AALEAKICHQI-EYFVGDF 28

RESULT 2

A31888

ribonucleoprotein La - human

N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B; Sjogren syndrome L

C:Species: Homo sapiens (man)

C>Date: 21-May-1990 #sequence_revision 26-May-1994 #text_change 22-Jun-1999

C:Accession: A31888; S03848; A22956; A61051; S11013; I55553; I70205; I70206; A312

R:Chambers, J.C.; Kenan, D.; Martin, B.J.; Keene, J.D.

J. Biol. Chem. 263, 18043-18051, 1988

A:Title: Genomic structure and amino acid sequence domains of the human La autoan

A:Reference number: A31888; MUID:89053970; PMID:3192525

A:Accession: A31888

A:Molecule type: mRNA

A:Residues: 1-408 <CHA>

A:Cross-references: GB:J04205; NID:g178686; PIDN:AAA51885.1; PID:g178687

R:Chan, E.K.L.; Sullivan, K.F.; Tan, E.M.

Nucleic Acids Res. 17, 2233-2244, 1989

A:Title: Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequ

A:Reference number: S03848; MUID:89202037; PMID:2460131
 A:Accession: S03848
 A:Molecule type: mRNA
 A:Residues: 1-408 <CH2>
 A:Cross-references: EMBL:X13697; NID:g36414; PIDN:CAA31985.1; PID:g36415
 R:Chambers, J.C.; Keene, U.S.A. 82, 2115-2119, 1985
 Proc. Natl. Acad. Sci. U.S.A.
 A:Title: Isolation and analysis of cDNA clones expressing human lupus La antigen.
 A:Reference number: A22956; MUID:85166283; PMID:3856888
 A:Accession: A22956
 A:Molecule type: mRNA
 A:Residues: 45-97, 'LK' <CH3>
 A:Cross-references: GB:J04205
 A:Note: this sequence has been revised in reference A31888
 R:Nyman, U.; Ringertz, N.R.; Pettersson, I.
 Immunol. Lett. 22, 65-72, 1989
 A:Title: Demonstration of an amino terminal La epitope recognized by human anti-La sera.
 A:Reference number: A61051; MUID:89379261; PMID:2476379
 A:Accession: A61051
 A>Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-19, 'E', '21-47 <NYM>
 R:Sturgess, A.D.; Peterson, M.D.; Miyamoto, K.; Fujii, H.; Miura, H.; Miyasaka, K.; Miyamoto, J. Clin. Invest. 85, 1566-1574, 1990
 A:Title: Fine epitope mapping the human SS-B/La protein: Identification of a distinct autoantigenic site.
 A:Reference number: I55553; MUID:90237237; PMID:1692037
 A:Accession: I55553
 A>Status: translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 81-107 <RES>
 A:Cross-references: GB:M35261; NID:g338491; PIDN:AAA36652.1; PID:g338495
 A:Accession: I70205
 A>Status: translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 174-224 <RE2>
 A:Cross-references: GB:M35263; NID:g338492; PIDN:AAA36653.1; PID:g338496
 A:Accession: I70206
 A>Status: translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 279-342 <RE3>
 A:Cross-references: GB:M35262; NID:g338493; PIDN:AAA36654.1; PID:g338497
 C:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
 C:Genetics:
 A:Gene: GDB:SSB
 A:Cross-references: GDB:I25359; OMIM:109090
 A:Map position: 2
 A:Introns: 22/3; 57/2; 115/3; 185/2; 209/2; 223/3; 264/3; 380/2
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:228-408/Domain: phosphorylated #status experimental <PHY>
 Query Match 85.18; Score 88.5; DB 1; Length 408;
 Best Local Similarity 94.7%; Pred. No. 3.3e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 Qy 1 AALEAKICHQIEEYFGDF 19
 Db 11 AALEAKICHQI-EYIFGDF 28
 RESULT 3
 S33818
 ribonucleoprotein La.A - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33817; S28544
 R:Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression
 A:Reference number: S33817; MUID:93287095; PMID:8510143
 A:Accession: S33817
 A:Molecule type: mRNA
 A:Residues: 1-427 <SCH>
 A:Cross-references: EMBL:X68818; NID:g64875; PIDN:CAA48716.1; PID:g64876
 C:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:111-177/Domain: ribonucleoprotein repeat homology <RRM>
 F:112-117/Region: RNA-binding RNP2 motif
 F:150-157/Region: RNA-binding RNP1 motif
 F:227-427/Domain: phosphorylated #status predicted <PHY>
 Query Match 57.28; Score 59.5; DB 1; Length 427;
 Best Local Similarity 75.08%; Pred. No. 0.13;
 Matches 12; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 Qy 3 LEAKICHQIEEYFGD 18
 Db 12 LDTKICEQI-EYIFGD 26
 RESULT 5
 S33818
 ribonucleoprotein La.A - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33818; S28545

JC1494
 ribonucleoprotein La - rat
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 24-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: JC1494; S25145
 R:Semsei, I.; Troester, H.; Bartsch, H.; Schwemmler, M.; Igloi, G.L.; Bachmann, M.
 Gene 126, 265-268, 1993
 A:Title: Isolation of rat cDNA clones coding for the autoantigen SS-B/La: Detection
 A:Reference number: JC1494; MUID:93246255; PMID:7916708
 A:Accession: JC1494
 A:Molecule type: mRNA
 A:Residues: 1-415 <SEM>
 A:Cross-references: GB:X67859; NID:g55778; PIDN:CAA48043.1; PID:g55779
 A:Experimental source: liver
 C:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:112-178/Domain: ribonucleoprotein repeat homology <RRM>
 F:113-118/Region: RNA-binding RNP2 motif
 F:151-158/Region: RNA-binding RNP1 motif
 F:227-415/Domain: phosphorylated #status predicted <PHY>
 Query Match 85.18; Score 88.5; DB 1; Length 415;
 Best Local Similarity 94.7%; Pred. No. 3.4e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 Qy 1 AALEAKICHQIEEYFGDF 19
 Db 11 AALEAKICHQI-EYIFGDF 28
 RESULT 4
 S33817
 ribonucleoprotein La.B - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33817; S28544
 R:Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.
 J. Mol. Biol. 231, 196-204, 1993
 A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expression
 A:Reference number: S33817; MUID:93287095; PMID:8510143
 A:Accession: S33817
 A:Molecule type: mRNA
 A:Residues: 1-427 <SCH>
 A:Cross-references: EMBL:X68818; NID:g64875; PIDN:CAA48716.1; PID:g64876
 C:Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.
 C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
 C:Keywords: phosphoprotein; RNA binding
 F:111-177/Domain: ribonucleoprotein repeat homology <RRM>
 F:112-117/Region: RNA-binding RNP2 motif
 F:150-157/Region: RNA-binding RNP1 motif
 F:227-427/Domain: phosphorylated #status predicted <PHY>
 Query Match 57.28; Score 59.5; DB 1; Length 427;
 Best Local Similarity 75.08%; Pred. No. 0.13;
 Matches 12; Conservative 1; Mismatches 2; Indels 1; Gaps 1;
 Qy 3 LEAKICHQIEEYFGD 18
 Db 12 LDTKICEQI-EYIFGD 26
 RESULT 5
 S33818
 ribonucleoprotein La.A - African clawed frog
 N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
 C:Accession: S33818; S28545

R;Schervly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.

J. Mol. Biol. 231, 196-204, 1993

A;Title: La proteins from *Xenopus laevis*. cDNA cloning and developmental expression.

A;Reference number: S33817; MUID:93287095; PMID:8510143

A;Accession: S33818

A;Molecule type: mRNA

A;Residues: 1-428 <SCH>

A;Cross-references: EMBL:X68817; NID:964873; PIDN:CAA48715.1; PID:964874

C;Comment: This protein associates with a variety of small RNA molecules, most of which act as a transcription termination factor.

C;Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology

C;Keywords: phosphoprotein; RNA binding

F;112-178/Domain: ribonucleoprotein repeat homology <RRM>

F;113-118/Region: RNA-binding RNP2 motif

F;151-158/Region: RNA-binding RNP1 motif

F;228-428/Domain: phosphorylated #status predicted <PHY>

Query Match 53.4%; Score 55.5; DB 1; Length 428;
Best Local Similarity 73.3%; Pred. No. 0.56;
Matches 11; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 4 EAKICHQIEEYFGD 18

Db 14 DTRICEQI-EYVFGD 27

RESULT 6

B70456

Thioredoxin - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999

C;Accession: B70456

R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ovi-

Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: B70456

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-135 <AQF>

A;Cross-references: GB:AE000757; NID:92984092; PIDN:AA07635.1; PID:92984097; GB:AE00065

A;Experimental source: strain VF5

C;Genetics:

A;Gene: trxA2

Query Match

Best Local Similarity 52.9%; Score 55; DB 2; Length 135;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 CHQIEEYFGD 18

Db 52 CHQVEEYFGD 62

RESULT 7

AC1870

hypothetical protein alr0508 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp.

A;Note: Nostoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002

C;Accession: AC1870

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana-*

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AC1870

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-187 <KUR>

A;Cross-references: GB:BA000019; PIDN:BA072466.1; PID:917129853; GSPDB:GN00179

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: alr0508

Query Match 45.2%; Score 47; DB 2; Length 187;

Best Local Similarity 50.0%; Pred. No. 5.5;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICHQIEEYFGD 18

Db 97 ICHELESWFLGD 108

RESULT 8

H90120

hypothetical protein orf357 [imported] - *Guillardia theta* nucleomorph

C;Species: nucleomorph *Guillardia theta*

A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001

C;Accession: H90120

R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X

Nature 410, 1091-1096, 2001

A;Title: The highly reduced genome of an enslaved algal nucleus.

A;Reference number: A9082; MUID:11323671; PMID:11323671

A;Accession: H90120

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-357 <DOU>

A;Cross-references: GB:AF083031; NID:913794284; PIDN:AAK39661.1; GSPDB:GN00152

C;Genetics:

A;Gene: orf357

A;Map position: 3

A;Genome: nucleomorph

C;Keywords: nucleomorph

Query Match 45.2%; Score 47; DB 2; Length 357;

Best Local Similarity 50.0%; Pred. No. 10;

Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 LEAKICHQIEEYFGD 18

Db 242 LKNKVFQMKNYFGD 257

RESULT 9

T39314

hypothetical protein SPBC119.16c - fission yeast (*Schizosaccharomyces pombe*)

C;Species: *Schizosaccharomyces pombe*

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C;Accession: T39314

R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.

submitted to the EMBL Data Library, March 1998

A;Reference number: Z21843

A;Accession: T39314

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-448 <WOO>

A;Cross-references: EMBL:AL022117; PIDN:CAA17931.1; GSPDB:GN00067; SPDB:SPEC119.1

A;Experimental source: strain 972h-; cosmid c119

C;Genetics:

A;Gene: SPDB:SPBC119.16c

A;Map position: 2

Query Match 45.2%; Score 47; DB 2; Length 448;

Best Local Similarity 47.4%; Pred. No. 13;

Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEEYFGD 19

Db 9 ATFDAKEGYNYYPGDF 27

RESULT 10

B35843

lipopolysaccharide-binding protein - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 05-Oct-1990 #sequence_revision 05-Oct-1990 #text_change 20-Aug-1999
 C:Accession: B35843; A46553
 R:Schumann, R.R.; Leong, S.R.; Flaggs, G.W.; Gray, P.W.; Wright, S.D.; Mathison, J.C.; T
 Science 249, 1429-1431, 1990
 A:Title: Structure and function of lipopolysaccharide binding protein.
 A:Reference number: A35843; MUID:90385281; PMID:2402637
 A:Accession: B35843
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-482 <SCH>
 A:Cross-references: GB:M35534; NID:g165467; PIDN:AAA99235.1; PID:g165468
 R:Tobias, P.S.; Soldau, K.; Ulevitch, R.J.
 J. Exp. Med. 164, 777-793, 1986
 A:Title: Isolation of a lipopolysaccharide-binding acute phase reactant from rabbit seru
 A:Reference number: A46553; MUID:86306528; PMID:2427635
 A:Accession: A46553
 A:Molecule type: protein
 A:Residues: 27-55, 'XG', 58-62, 'F', 64-65 <TOB>
 C:Superfamily: lipopolysaccharide-binding protein
 C:Keywords: acute phase

Query Match 45.2%; Score 47; DB 2; Length 482;
 Best Local Similarity 81.8%; Pred. No. 14;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 3 LEAKICHQIEE 13
 ||||| |||||
 Db 194 LESKICRQIEE 204

RESULT 11
 S48331
 Probable membrane protein YML067c - yeast (Saccharomyces cerevisiae)
 C:Species: Saccharomyces cerevisiae
 C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 19-Apr-2002
 C:Accession: S48331; S48330
 R:Bowman, S.; Churcher, C.
 submitted to the EMBL Data Library, August 1994
 A:Reference number: S48326
 A:Accession: S48331
 A:Molecule type: DNA
 A:Residues: 1-352 <BO>
 A:Cross-references: EMBL:Z38114; NID:g558402; PID:g558408; GSPDB:GN00013; MIPS:YML067c
 A:Accession: S48330
 A:Residues: 69-352 <BO>
 A:Cross-references: EMBL:Z38114; NID:g558402; PID:g558407
 C:Genetics:
 A:Gene: SGD:ERV41; MIPS:YML067c
 A:Cross-references: SGD:S0004532
 A:Map position: 13L
 A:Introns: 11/1
 C:Keywords: transmembrane protein
 F:230-245/Domain: transmembrane #status predicted <TM1>
 F:248-264/Domain: transmembrane #status predicted <TM2>

Query Match 44.2%; Score 46; DB 2; Length 352;
 Best Local Similarity 56.2%; Pred. No. 14;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 4 EAKICHQIEEYFGDF 19
 ||||| |||||
 Db 196 ELKFHNEVINEFSFGDF 211

RESULT 12
 E69391
 hypothetical protein AF1134 - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
 C:Accession: E69391

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; D
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness,
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Syke
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing ar
 A:Reference number: A69250; MUID:98049343; PMID:9389475
 A:Accession: E69391
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-456 <KLE>
 A:Cross-references: GB:AE001026; GB:AE000782; NID:g2689349; PIDN:AA890123.1; PID:g

Query Match 44.2%; Score 46; DB 2; Length 456;
 Best Local Similarity 61.5%; Pred. No. 19;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 3 LEAKICHQIEEY 15
 :||| |||||
 Db 405 MEAKTPHOFDEY 417

RESULT 13
 H30096
 eukaryotic release factor 1 homolog [Imported] - Guillardia theta nucleomorph
 C:Species: nucleomorph Guillardia theta
 A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 15-Jun-2001
 C:Accession: H30096
 R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X
 Nature 410, 1091-1096, 2001
 A:Title: The highly reduced
 A:Reference number: A39082; MUID:11323671; PMID:11323671
 A:Accession: H30096
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-409 <DOU>
 A:Cross-references: GB:AF165818; NID:g13794528; PIDN:AAK39903.1; GSPDB:GN00150
 C:Genetics:
 A:Gene: erfl
 A:Map position: 1
 A:Genome: nucleomorph
 C:Superfamily: translation releasing factor erf-1
 C:Keywords: nucleomorph

Query Match 43.3%; Score 45; DB 2; Length 409;
 Best Local Similarity 50.0%; Pred. No. 24;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 6 KICHOIEEYFGDF 19
 ||||| :||| ||
 Db 200 KICELADQYILSDF 213

RESULT 14
 A54413
 acetylcholinesterase (EC 3.1.1.7) A precursor - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 21-Jul-2000
 C:Accession: A54413; T29824
 R:Arpagaus, M.; Fedon, Y.; Cousin, X.; Chatoñnet, A.; Berge, J.B.; Fournier, D.; T
 J. Biol. Chem. 269, 9957-9965, 1994
 A:Title: cDNA sequence, gene structure, and in vitro expression of ace-1, the gene
 A:Reference number: A54413; MUID:94193691; PMID:8144590
 A:Accession: A54413
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-620 <ARP>
 A:Cross-references: GB:X75331; NID:g475060; PIDN:CAA53080.1; PID:g671831
 R:Wu, X.; Le, T.T.
 submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of C. elegans cosmid W09B12.

A;Reference number: Z20693
A;Accession: T29824
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-620 <WUX>
A;Cross-references: EMBL:U58731; PIDN:AAB00593.1; GSPDB:GN00028
A;Experimental source: strain Bristol N2; clone W09B12
C;Genetics:
A;Gene: CESP:ace-1
A;Map position: X
A;Introns: 13/3; 59/2; 154/2; 236/3; 296/2; 454/2; 509/2; 573/1; 606/2
C;Superfamily: cholinesterase; cholinesterase homology
C;Keywords: carboxylic ester hydrolase
F;45-567/Domain: cholinesterase homology <CHE>

Query Match 43.3%; Score 45; DB 2; Length 620;
Best Local Similarity 66.7%; Pred. No. 36;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 CHOIEEYFGDF 19
| | | | |
Db 82 CIOSEDYFGDF 93

RESULT 15

T37411
RNA polymerase subunit rpol32 - vaccinia virus (strain Ankara)
C;Species: vaccinia virus
A;Variety: strain Ankara
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
C;Accession: T37411
R;Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dorner, F.
submitted to the EMBL Data Library, March 1997
A;Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
A;Reference number: Z20877
A;Accession: T37411
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1156 <ANT>
A;Cross-references: EMBL:U94848; PIDN:AAB96526.1
A;Experimental source: strain Ankara
C;Genetics:
A;Note: MVA135R
C;Superfamily: DNA-directed RNA polymerase 132K polypeptide

Query Match 43.3%; Score 45; DB 2; Length 1156;
Best Local Similarity 56.2%; Pred. No. 65;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 LEAKICHOIEEYFGD 18
| | | | |
Db 497 LEKKICYIRSYKDD 512

Search completed: April 23, 2003, 13:34:40
Job time : 12.6742 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:27:28 ; Search time 10.8876 Seconds
(without alignments)
139.834 Million cell updates/sec

Title: US-09-836-073-16

Perfect score: 104

Sequence: 1 AALEAKICHQIEEYFGDF 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2.6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
- 2: /cgn2.6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
- 3: /cgn2.6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
- 4: /cgn2.6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
- 5: /cgn2.6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
- 6: /cgn2.6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
- 7: /cgn2.6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
- 8: /cgn2.6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
- 9: /cgn2.6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
- 10: /cgn2.6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
- 11: /cgn2.6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
- 12: /cgn2.6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
- 13: /cgn2.6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
- 14: /cgn2.6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	19	US-09-836-073-16	Sequence 16, Appl
2	88.5	85.1	18	US-09-836-073-1	Sequence 1, Appl
3	88.5	85.1	18	US-09-836-073-14	Sequence 14, Appl
4	88.5	85.1	460	US-10-102-806-695	Sequence 695, Appl
5	84.5	81.2	17	US-09-836-073-13	Sequence 13, Appl
6	83.5	80.3	18	US-09-836-073-11	Sequence 11, Appl
7	82.5	79.3	18	US-09-836-073-9	Sequence 9, Appl
8	80.5	77.4	18	US-09-836-073-12	Sequence 12, Appl
9	79.5	76.4	18	US-09-836-073-10	Sequence 10, Appl
10	76.5	73.6	18	US-09-836-073-2	Sequence 2, Appl
11	76.5	73.6	18	US-09-836-073-3	Sequence 3, Appl
12	76.5	73.6	18	US-09-836-073-4	Sequence 4, Appl
13	75.5	72.6	18	US-09-836-073-8	Sequence 8, Appl
14	73.5	70.7	18	US-09-836-073-7	Sequence 7, Appl
15	65.5	63.0	18	US-09-836-073-15	Sequence 15, Appl
16	64.5	62.0	18	US-09-836-073-5	Sequence 5, Appl
17	63	60.6	18	US-09-836-073-6	Sequence 6, Appl
18	58	55.8	37	US-09-843-676-24	Sequence 24, Appl
19	58	55.8	37	US-09-766-253-24	Sequence 24, Appl

20	55.8	37	9	US-09-438-486-24	Sequence 24, Appl
21	55.8	37	9	US-10-053-758-24	Sequence 24, Appl
22	55.8	37	9	US-10-054-295-24	Sequence 24, Appl
23	55.8	37	9	US-10-054-611-24	Sequence 24, Appl
24	49	2169	9	US-09-738-626-5455	Sequence 5455, Appl
25	48.5	38	9	US-09-843-676-25	Sequence 25, Appl
26	48.5	38	9	US-09-766-253-25	Sequence 25, Appl
27	48.5	38	9	US-09-438-486-25	Sequence 25, Appl
28	48.5	38	9	US-10-053-758-25	Sequence 25, Appl
29	48.5	38	9	US-10-054-295-25	Sequence 25, Appl
30	48.5	38	9	US-10-054-611-25	Sequence 25, Appl
31	47	456	10	US-09-861-400-6	Sequence 6, Appl
32	47	482	10	US-09-861-400-5	Sequence 5, Appl
33	46.5	16	9	US-09-836-073-19	Sequence 19, Appl
34	45.5	643	9	US-09-736-968A-60	Sequence 60, Appl
35	45.5	2045	9	US-09-736-968A-109	Sequence 109, Appl
36	45.5	2045	10	US-09-736-969A-95	Sequence 95, Appl
37	45.5	2045	10	US-09-736-960-92	Sequence 92, Appl
38	45.5	2047	9	US-09-736-968A-2	Sequence 2, Appl
39	43	23	9	US-10-011-585A-144	Sequence 144, Appl
40	43	88	10	US-09-955-502-26	Sequence 26, Appl
41	43	88	10	US-09-955-502-27	Sequence 27, Appl
42	43	88	10	US-09-955-502-28	Sequence 28, Appl
43	42	18	9	US-09-836-073-18	Sequence 18, Appl
44	42	186	9	US-10-013-315-6	Sequence 6, Appl
45	41	184	10	US-09-734-017A-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; TYPE: PRT
; LENGTH: 19
; ORGANISM: Rat
US-09-836-073-16

Query Match 100.0%; Score 104; DB 9; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.1e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALEAKICHQIEEYFGDF 19
|||||

Db 1 AALEAKICHQIEEYFGDF 19

RESULT 2
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match 85.1%; Score 88.5; DB 9; Length 18;
Best Local Similarity 94.7%; Pred. No. 8.8e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
|||||
Db 1 AALEAKICHQI-EYYFGDF 18

RESULT 3

US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14

Query Match 85.1%; Score 88.5; DB 9; Length 18;
Best Local Similarity 94.7%; Pred. No. 8.8e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
|||||
Db 1 AALEAKICHQI-EYYFGDF 18

RESULT 4

US-10-102-806-695
; Sequence 695, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 695
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-102-806-695

Query Match 85.1%; Score 88.5; DB 9; Length 460;
Best Local Similarity 94.7%; Pred. No. 2.3e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
|||||
Db 63 AALEAKICHQI-EYYFGDF 80

RESULT 5

US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match 81.2%; Score 84.5; DB 9; Length 17;
Best Local Similarity 94.4%; Pred. No. 3.6e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2 AALEAKICHQIEYYFGDF 19
|||||
Db 1 AALEAKICHQI-EYYFGDF 17

RESULT 6

US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match 80.3%; Score 83.5; DB 9; Length 18;
Best Local Similarity 89.5%; Pred. No. 5.5e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
|||||
Db 1 AALEAKICHQIEQ-YFGDF 18

```

RESULT 7
US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9
Query Match      79.3%; Score 82.5; DB 9; Length 18;
Best Local Similarity 94.4%; Pred. No. 8e-07; 0; Indels 1; Gaps 1;
Matches 17; Conservative 0; Mismatches 1;

QY 1 AALEAKICHQIEEYFGD 18
Db 1 AALEAKICHQI-EYFGD 17

RESULT 8
US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12
Query Match      77.4%; Score 80.5; DB 9; Length 18;
Best Local Similarity 89.5%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEEYFGD 19
Db 1 AALEAKICHQI-EYFGD 18

RESULT 9
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

```

```

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10
Query Match      76.4%; Score 79.5; DB 9; Length 18;
Best Local Similarity 89.5%; Pred. No. 2.4e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEEYFGD 19
Db 1 AALEAKICHQI-EYFGD 18

RESULT 10
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2
Query Match      73.6%; Score 76.5; DB 9; Length 18;
Best Local Similarity 84.2%; Pred. No. 7.2e-06;
Matches 16; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEEYFGD 19
Db 1 AALEAKICHQI-EYFGD 18

RESULT 11
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

```

Query Match 73.6%; Score 76.5; DB 9; Length 18;
Best Local Similarity 78.9%; Pred. No. 7.2e-06;
Matches 15; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
| | | | | | | | | | | | | | | | | | | |
DB 1 AALQAKICHQI-QYFQGF 18

RESULT 12

US-09-836-073-4
Sequence 4, Application US/09836073

Patent No. US20020173475A1

GENERAL INFORMATION:

APPLICANT: Dasgupta, Asim

APPLICANT: Das, S.

APPLICANT: Balidya, Narayan

TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

FILE REFERENCE: 22002054822

CURRENT APPLICATION NUMBER: US/09/836,073

CURRENT FILING DATE: 2002-10-24

PRIOR APPLICATION NUMBER: 09/316,630

PRIOR FILING DATE: 1999-05-21

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 4

LENGTH: 18

TYPE: PRT

ORGANISM: Homo Sapiens

US-09-836-073-4

Query Match 73.6%; Score 76.5; DB 9; Length 18;
Best Local Similarity 93.8%; Pred. No. 7.2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 4 EAKICHQIEYYFGDF 19
| | | | | | | | | | | | | | | | | | | |
DB 4 EAKICHQI-EYFQGF 18

RESULT 13

US-09-836-073-8

Sequence 8, Application US/09836073

Patent No. US20020173475A1

GENERAL INFORMATION:

APPLICANT: Dasgupta, Asim

APPLICANT: Das, S.

APPLICANT: Balidya, Narayan

TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

FILE REFERENCE: 22002054822

CURRENT APPLICATION NUMBER: US/09/836,073

CURRENT FILING DATE: 2002-10-24

PRIOR APPLICATION NUMBER: 09/316,630

PRIOR FILING DATE: 1999-05-21

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 8

LENGTH: 18

TYPE: PRT

ORGANISM: Homo Sapiens

US-09-836-073-8

Query Match 72.6%; Score 75.5; DB 9; Length 18;
Best Local Similarity 84.2%; Pred. No. 1e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
| | | | | | | | | | | | | | | | | | | |
DB 1 AALEAKICHQIEQ-QFGDF 18

RESULT 14

US-09-836-073-7

Sequence 7, Application US/09836073
Patent No. US20020173475A1
GENERAL INFORMATION:
APPLICANT: Dasgupta, Asim
APPLICANT: Das, S.
APPLICANT: Balidya, Narayan

TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

FILE REFERENCE: 22002054822

CURRENT APPLICATION NUMBER: US/09/836,073

CURRENT FILING DATE: 2002-10-24

PRIOR APPLICATION NUMBER: 09/316,630

PRIOR FILING DATE: 1999-05-21

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 7

LENGTH: 18

TYPE: PRT

ORGANISM: Homo Sapiens

US-09-836-073-7

Query Match 70.7%; Score 73.5; DB 9; Length 18;
Best Local Similarity 88.9%; Pred. No. 2.2e-05;
Matches 16; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 18
| | | | | | | | | | | | | | | | | | | |
DB 1 AALEAKICHQI-EYQGD 17

RESULT 15

US-09-836-073-15

Sequence 15, Application US/09836073

Patent No. US20020173475A1

GENERAL INFORMATION:

APPLICANT: Dasgupta, Asim

APPLICANT: Das, S.

APPLICANT: Balidya, Narayan

TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

FILE REFERENCE: 22002054822

CURRENT APPLICATION NUMBER: US/09/836,073

CURRENT FILING DATE: 2002-10-24

PRIOR APPLICATION NUMBER: 09/316,630

PRIOR FILING DATE: 1999-05-21

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 15

LENGTH: 18

TYPE: PRT

ORGANISM: Xenopus

US-09-836-073-15

Query Match 63.0%; Score 65.5; DB 9; Length 18;
Best Local Similarity 76.5%; Pred. No. 0.00041;
Matches 13; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 3 LEAKICHQIEYYFGDF 19
| | | | | | | | | | | | | | | | | | | |
DB 3 LDTKICEQI-EYFQGF 18

Search completed: April 23, 2003, 13:38:21
Job time : 11.8876 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:24 ; Search time 11.8483 Seconds
(without alignments)
47.183 Million cell updates/sec

Title: US-09-836-073-16
Perfect score: 104
Sequence: 1 AALEAKICHQIEEYFGDF 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2.6/prodata/1/1aa/5A_COMB.pep.*
2: /cgn2.6/prodata/1/1aa/5B_COMB.pep.*
3: /cgn2.6/prodata/1/1aa/6A_COMB.pep.*
4: /cgn2.6/prodata/1/1aa/6B_COMB.pep.*
5: /cgn2.6/prodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2.6/prodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88.5	85.1	18	4	US-09-316-630-3
2	88.5	85.1	18	4	US-09-316-630-4
3	58	55.8	37	3	US-08-851-843A-24
4	58	55.8	37	4	US-08-854-050-24
5	58	55.8	37	4	US-09-430-323-24
6	57.5	55.3	38	4	US-08-974-549A-214
7	48.5	46.6	38	3	US-08-851-843A-25
8	48.5	46.6	38	4	US-08-974-549A-215
9	48.5	46.6	38	4	US-08-854-050-25
10	48.5	46.6	38	4	US-09-430-323-25
11	47	45.2	456	1	US-08-205-719-4
12	47	45.2	456	4	US-08-431-517F-6
13	47	45.2	482	4	US-08-431-517F-5
14	47	45.2	482	6	5245013-2
15	43	41.3	510	1	US-08-097-829-2
16	43	41.3	510	1	US-08-577-403-2
17	42.5	40.9	740	4	US-09-323-872A-23
18	42.5	40.9	864	4	US-09-323-872A-28
19	40	38.5	263	2	US-08-864-799-1
20	40	38.5	1253	1	US-07-920-281C-3
21	40	38.5	1253	4	US-08-466-277-3
22	39	37.5	86	4	US-09-300-008B-55
23	39	37.5	282	3	US-08-972-902-7
24	39	37.5	811	4	US-09-199-637A-93
25	38	36.5	461	2	US-08-672-814D-2
26	38	36.5	461	4	US-09-333-696-2
27	38	36.5	488	4	US-08-910-313-2

28	38	36.5	984	4	US-09-287-354-2	Sequence 2, Appl1
29	38	36.5	1189	4	US-09-287-354-3	Sequence 3, Appl1
30	38	36.5	1189	4	US-09-287-354-4	Sequence 4, Appl1
31	38	36.5	1207	4	US-09-287-354-5	Sequence 5, Appl1
32	37.5	36.1	39	3	US-08-851-843A-26	Sequence 26, Appl1
33	37.5	36.1	39	4	US-08-974-549A-216	Sequence 216, Appl
34	37.5	36.1	39	4	US-08-854-050-26	Sequence 26, Appl
35	37.5	36.1	39	4	US-09-430-323-26	Sequence 26, Appl
36	37	35.6	346	2	US-08-744-779A-2	Sequence 2, Appl1
37	37	35.6	346	2	US-08-862-531-2	Sequence 2, Appl1
38	37	35.6	432	4	US-09-282-305-10	Sequence 10, Appl1
39	37	35.6	478	4	US-09-326-529-2	Sequence 2, Appl1
40	37	35.6	573	4	US-09-330-740A-8	Sequence 8, Appl1
41	37	35.6	748	1	US-08-408-318-2	Sequence 2, Appl1
42	37	35.6	748	1	US-08-369-796-10	Sequence 10, Appl
43	37	35.6	748	1	US-08-839-164-2	Sequence 2, Appl1
44	37	35.6	748	2	US-08-852-091-10	Sequence 10, Appl
45	37	35.6	748	2	US-08-820-754-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match 85.1%; Score 88.5; DB 4; Length 18;
Best Local Similarity 94.7%; Pred. No. 3.7e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEEYFGDF 19
|||||
Db 1 AALEAKICHQI-EYFGDF 18

RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
; OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 85.1%; Score 88.5; DB 4; Length 18;
Best Local Similarity 94.7%; Pred. No. 3.7e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 AALEAKICHQIEEYFGDF 19
| | | | | | | | | | | | | | | | | |
Db 1 AALEAKICHQI-EYFVGDF 18

RESULT 3
US-08-851-843A-24
; Sequence 24, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 24:

; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-851-843A-24

Query Match 55.8%; Score 58; DB 3; Length 37;
Best Local Similarity 84.6%; Pred. No. 0.0071;
Matches 11; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

Qy 7 ICHQIEEYFGDF 19
| | | | | | | | | | | | | | | | | |
Db 1 ICHQ-EYFVGDF 11

RESULT 4
US-08-854-050-24
; Sequence 24, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-854-050-24

Query Match 55.88; Score 58; DB 4; Length 37;
Best Local Similarity 84.68; Pred. No. 0.0071;
Matches 11; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 7 ICHOIEEYFGDF 19
||||| |||||
Db 1 ICHQ--EYFGDF 11

RESULT 5
US-09-430-323-24
; Sequence 24, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; LINGNER, Joachim
; NAKAMURA, Toru
; CHAPMAN, Karen B.
; MORIN, Gregg B.
; HARLEY, Calvin B.
; ANDREWS, William H.

; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430.323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996

; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:

US-09-430-323-24

Query Match 55.88; Score 58; DB 4; Length 37;
Best Local Similarity 84.68; Pred. No. 0.0071;
Matches 11; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 7 ICHOIEEYFGDF 19
||||| |||||
Db 1 ICHQ--EYFGDF 11

RESULT 6
US-08-974-549A-214
; Sequence 214, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17618
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph Ted

REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-214

Query Match 55.3%; Score 57.5; DB 4; Length 38;
Best Local Similarity 84.6%; Pred. No. 0.0088;
Matches 11; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIEYYFGDF 19
||||| |||||
Db 1 ICHQ-XEYFGDF 12

RESULT 7
US-08-851-843A-25
Sequence 25, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morlin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
CLASSIFICATION:
PRIOR APPLICATION NUMBER: WO PCT/US97/17618

SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-25

Query Match 46.6%; Score 48.5; DB 3; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.25;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIEYYFGDF 18
||||| |||||
Db 1 ICEQI-EYFGD 11

RESULT 8
US-08-974-549A-215
Sequence 215, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morlin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618

```
;
; FILING DATE: 01-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 215:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-974-549A-215

Query Match 46.6%; Score 48.5; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.25;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIEYYFGD 18
Db 1 ICEQI-EYYFGD 11

RESULT 9
US-08-854-050-25
; Sequence 25, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
```

```
;
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-854-050-25

Query Match 46.6%; Score 48.5; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.25;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICHQIEYYFGD 18
Db 1 ICEQI-EYYFGD 11

RESULT 10
US-09-430-323-25
; Sequence 25, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
```

; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match 45.6%; Score 48.5; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.25;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 7 ICHQIEYYFGD 18
||| |||||
DB 1 ICEQI-EYVFGD 11

RESULT 11

US-08-205-719-4
; Sequence 4, Application US/08205719
; Patent No. 5705398

; GENERAL INFORMATION:
; APPLICANT: Mintz, D. N.
; APPLICANT: Tobias, P. S.
; APPLICANT: Ulevitch, R. J.
; TITLE OF INVENTION: SYSTEM AND METHODS FOR IDENTIFYING
; REFERENCE/DOCKET NUMBER: THERAPEUTIC INHIBITORS OF LPS-MEDIATED SEPSIS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5705398th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/205.719
; FILING DATE: 02-MAR-1994

; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1201P; TSRI324.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-205-719-4

Query Match 45.2%; Score 47; DB 1; Length 456;
Best Local Similarity 81.8%; Pred. No. 6.4;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 3 LEAKICHOIEE 13
||:|||||
DB 168 LESKICRQIEE 178

RESULT 12
US-08-431-517F-6
; Sequence 6, Application US/08431517F
; Patent No. 6265187

; GENERAL INFORMATION:
; APPLICANT: Scott, Randal W
; APPLICANT: Marra, Marian N
; TITLE OF INVENTION: RECOMBINANT ENDOTOXIN-NEUTRALIZING PROTEINS
; FILE REFERENCE: 1103/11307US01
; CURRENT APPLICATION NUMBER: US/08/431.517F
; CURRENT FILING DATE: 1995-05-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 456

; TYPE: PRT
; ORGANISM: rabbit

; FEATURE:
; OTHER INFORMATION: rabbit LBP amino acid (Figure 5); mature protein
; OTHER INFORMATION: sequence
US-08-431-517F-6

Query Match 45.2%; Score 47; DB 4; Length 456;
Best Local Similarity 81.8%; Pred. No. 6.4;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 3 LEAKICHOIEE 13
||:|||||
DB 168 LESKICRQIEE 178

RESULT 13

US-08-431-517F-5
; Sequence 5, Application US/08431517F
; Patent No. 6265187

; GENERAL INFORMATION:
; APPLICANT: Scott, Randal W
; APPLICANT: Marra, Marian N
; TITLE OF INVENTION: RECOMBINANT ENDOTOXIN-NEUTRALIZING PROTEINS
; FILE REFERENCE: 1103/11307US01
; CURRENT APPLICATION NUMBER: US/08/431.517F
; CURRENT FILING DATE: 1995-05-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 482

; TYPE: PRT
; ORGANISM: rabbit

; FEATURE:
; OTHER INFORMATION: rabbit LBP amino acid (Figure 5)
; NAME/KEY: SIGNAL
; LOCATION: (1)..(26)
; NAME/KEY: CHAIN
; LOCATION: (27)..(482)
US-08-431-517F-5

Query Match 45.2%; Score 47; DB 4; Length 482;
Best Local Similarity 81.8%; Pred. No. 6.8;
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 3 LEAKICHOIEE 13
||:|||||
DB 194 LESKICRQIEE 204

RESULT 14

US-08-524-5013-2
; Patent No. 5245013

; APPLICANT: Ulevitch, Richard; Tobias, Peter
; TITLE OF INVENTION: ACUTE PHASE PROTEIN MODULATING ENDOTOXIC
; ACTIVITY OF LIPOPOLYSACCHARIDES, ASSAY METHODS AND POLYPEPTIDES
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/367,454
 ; FILING DATE: 01-JUN-1989
 ; APPLICATION NUMBER: 6,710
 ; FILING DATE: 30-DEC-1986
 ; APPLICATION NUMBER: 728,833
 ; FILING DATE: 30-APR-1985
 ; SEQ ID NO:2
 ; LENGTH: 482
 5245013-2

Query Match 45.2%; Score 47; DB 6; Length 482;
 Best Local Similarity 81.8%; Pred. No. 6.8;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LEAKICQIEE 13
 Db 194 LESKICQIEE 204

RESULT 15

US-08-097-829-2
 ; Sequence 2, Application US/08097829
 ; Patent No. 5498831
 ; GENERAL INFORMATION:
 ; APPLICANT: Burgess, Diane G.
 ; APPLICANT: Dooner, Hugo K.
 ; TITLE OF INVENTION: Pea ADP-Glucose Pyrophosphorylase Subunit
 ; TITLE OF INVENTION: Genes and Their Uses
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend Kourie and Crew
 ; STREET: One Market Plaza, Steuart Tower, Suite 2000
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/097,829
 ; FILING DATE: 19930723
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Heslin, James M.
 ; REGISTRATION NUMBER: 29,541
 ; REFERENCE/DOCKET NUMBER: 12176-35
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-326-2400
 ; TELEFAX: 415-326-2422
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 510 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-097-829-2

Query Match 41.3%; Score 43; DB 1; Length 510;
 Best Local Similarity 45.5%; Pred. No. 32;
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 9 HQIEEYFGDF 19
 Db 327 HNVQAYFFGDY 337

Search completed: April 23, 2003, 13:36:33
 Job time : 11.8483 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:23 ; Search time 30.1011 Seconds
(without alignments)
84.109 Million cell updates/sec

Title: US-09-836-073-16

Perfect score: 104

Sequence: 1 AALEAKICHQIEEYFGDF 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88.5	85.1	18	21	AAV52200 Human la autoantigen
2	88.5	85.1	92	21	AG01351 Human secreted pro
3	88.5	85.1	408	17	AAW03716 Human autoantigen
4	88.5	85.1	439	22	ABG08417 Novel human diagno
5	88.5	85.1	460	23	ABP1511 Breast and ovarian
6	88.5	85.1	460	23	ABP1511 Human ovarian anti
7	62.5	60.1	21	14	AA43394 La/SSB epitope 17
8	49	47.1	2159	22	AA43394 C glutamic prote
9	47	45.2	456	19	AAW41695 Rabbit LPS-binding
10	47	45.2	482	14	AA41659 Lapine gram-negati

11	47	45.2	482	15	AA33461 Lapine polysacchar
12	47	45.2	482	19	AAW40813 Rabbit lipopolysac
13	47	45.2	482	20	AAW78355 Rabbit lipopolysac
14	45.5	43.8	132	21	AA31905 Human secreted pro
15	45.5	43.8	425	22	AAW25631 Human protein sequ
16	45.5	43.8	643	22	AA39540 Preliminary human
17	45.5	43.8	2047	22	AA39541 Human CLASP-7 prot
18	45	43.3	379	21	AA35807 Arabidopsis thalia
19	45	43.3	387	21	AA35806 Arabidopsis thalia
20	45	43.3	467	21	AA35805 Arabidopsis thalia
21	44	42.3	176	22	AA52458 Protonibacterium
22	44	42.3	251	23	ABB76194 Human kidney injur
23	44	42.3	281	23	ABB76195 Rat kidney injur
24	44	42.3	284	23	ABB76193 Human kidney injur
25	44	42.3	592	22	ABB64658 Drosophila melanog
26	43	41.3	23	23	ABB79195 Human prostate spe
27	43	41.3	58	22	AAW4298 Human immune/haema
28	43	41.3	297	21	AA37294 Lipid associated p
29	43	41.3	297	22	AAW38768 Human polypeptide
30	43	41.3	297	22	AAW60090 Human transport pr
31	43	41.3	304	22	AAW40554 Human polypeptide
32	43	41.3	447	22	AAU00461 Arabidopsis thalia
33	43	41.3	510	17	AA394187 Pea ADPG-PPase SH2
34	43	41.3	510	19	AAW59447 Pea ADP-glucose py
35	42.5	40.9	740	20	AAW83066 Pyruvate formate 1
36	42.5	40.9	740	22	AAW59809 TtD protein #1.
37	42.5	40.9	864	20	AAW30576 Pyruvate formate 1
38	42.5	40.9	864	22	AAW59810 TtD protein #2.
39	42.5	40.9	1029	22	AAW59812 TtD protein #3.
40	42.5	40.9	1614	22	AAW59825 Protein #2 encoded
41	42	40.4	122	23	AAW10527 Human ORFX protein
42	42	40.4	175	18	AAW20649 H. pylori secreted
43	42	40.4	324	19	AAW98640 H. pylori GHP 346
44	42	40.4	324	19	AAW71525 Helicobacter polyp
45	42	40.4	386	23	ABB55321 Lactococcus lactis

ALIGNMENTS

RESULT 1
AAV52200
ID AAV52200 standard; peptide; 18 AA.
XX
AC AAV52200;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human la autoantigen peptide (LAP).

XX
KW La autoantigen; LAP; internal ribosome entry site; IRES; translation;
KW viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
KW coronavirus; hepatitis virus; rhinovirus; adenovirus; coxsackie virus;
KW parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
KW vesicular stomatitis virus.

OS Homo sapiens.

XX

PN WO9961613-A2.

XX

PD 02-DEC-1999.

XX

PF 21-MAY-1999; 99WO-US11281.

XX

PR 22-MAY-1998; 98US-0086527.

XX (REGC) UNIV CALIFORNIA.

PI Das S, Dasgupta A;

XX WPI; 2000-062712/05.

PT New yeast inhibitory peptide useful for inhibiting viral protein
 translation and replication -
 XX
 PS Claim 5; Page 57; 81pp; English.
 XX
 CC This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (IRNA see AAZ45200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rhinovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.
 XX
 SQ Sequence 18 AA;

Query Match 85.18; Score 88.5; DB 21; Length 18;
 Best Local Similarity 94.7%; Pred. No. 2.4e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 AALEAKICHQIEEYFGDF 19
 |||||
 Db 1 AALEAKICHQI-EYFGDF 18

RESULT 2
 AAG01351
 ID AAG01351 standard; Protein; 92 AA.
 XX
 AC AAG01351;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 5432.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45..
 DR N-PSDB; AAC01357.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13; SEQ ID 5432; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30

CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SQ Sequence 92 AA;

Query Match 85.18; Score 88.5; DB 21; Length 92;
 Best Local Similarity 94.7%; Pred. No. 1.4e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 AALEAKICHQIEEYFGDF 19
 |||||
 Db 11 AALEAKICHQI-EYFGDF 28

RESULT 3
 AAW03716
 ID AAW03716 standard; protein; 408 AA.
 XX
 AC AAW03716;
 XX
 DT 12-MAR-1997 (first entry)
 XX
 DE Human autoantigen La(SS-B).
 XX
 KW Autoimmune disease; La autoantigen; Sjogren's syndrome;
 KW systemic lupus erythematosus; diagnosis.
 XX
 OS Homo sapiens.
 XX
 PN US5541291-A.
 XX
 PD 30-JUL-1996.
 XX
 PF 31-DEC-1984; 84US-0687908.
 XX
 PR 27-MAY-1987; 87US-0054871.
 PR 31-DEC-1984; 84US-0687908.
 XX
 PA (UYDU-) UNIV DUKE.
 XX
 PI Keene JD;
 XX
 DR WPI; 1996-362015/36.
 XX
 PT Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
 PT overlap syndrome - useful for diagnosis and treatment of autoimmune
 PT diseases
 XX
 PS Disclosure; Columns 15-16; 21pp; English.
 XX
 CC The human lupus antigen (La) is diagnostic for Sjogren's syndrome,
 CC as well as occurring in systemic lupus erythematosus patients. The
 CC La protein is clinically related to the Ro protein that is highly
 CC common among autoimmune patients. La and Ro antigens sometimes
 CC reside on the same cellular ribonucleoprotein particle; most La
 CC patients contain some Ro antibodies and vice versa. La cDNA has
 CC been isolated from a human liver library.
 XX
 SQ Sequence 408 AA;

Query Match 85.18; Score 88.5; DB 17; Length 408;
 Best Local Similarity 94.7%; Pred. No. 7e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 AALEAKICHQIEEYFGDF 19

Db 11 AALEAKICHQI-EYFGDF 28
|||||

RESULT 4
ABG08417
ID ABG08417 standard; Protein; 439 AA.
XX
AC ABG08417;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #8408.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
PI WPI; 2001-639362/73.
XX
DR N-PSDB; AAS72604.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 38776; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 439 AA;

Query Match 85.1%; Score 88.5; DB 22; Length 439;
Best Local Similarity 94.7%; Pred. No. 7.6e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYFGDF 19
|||||
Db 41 AALEAKICHQI-EYFGDF 58

RESULT 5
AAB58987
ID AAB58987 standard; Protein; 460 AA.
XX
AC AAB58987;
XX
DT 27-MAR-2001 (first entry)
XX
DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 695.
XX
KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neoplastic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; vulnarary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.
XX
OS Homo sapiens.
XX
PN WO200055173-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05881.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
PI WPI; 2000-611515/58.
XX
DR N-PSDB; AAF21890.
XX
PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases -
XX
PS Claim 11; Page 1149-11150; 1299pp; English.
XX
CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC neoplastic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antitumor; vulnarary; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiant activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.
XX
SQ Sequence 460 AA;

Query Match 85.1%; Score 88.5; DB 21; Length 460;
Best Local Similarity 94.7%; Pred. No. 8e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYFGDF 19
|||||

Db 63 AALEAKICHQI-EYFGDF 80

RESULT 6

ABP41511

ID ABP41511 standard; Protein; 460 AA.

XX

AC ABP41511;

XX

DT 22-AUG-2002 (first entry)

XX

DE Human ovarian antigen HVVAF56, SEQ ID NO:2643.

XX

KW Human: ovarian antigen; ovary; ovarian; breast; cancer; tumour;

KW ovarian cancer; breast cancer; tumour; reproductive system disorder;

KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;

KW inflammatory condition; immune disorder; blood disorder;

KW cardiovascular disorder; respiratory disorder; neurological disorder;

KW gastrointestinal disorder; urinary system disorder; drug screening;

KW gene therapy; chromosome mapping; forensic analysis;

KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;

KW antiinflammatory; gynaecological; reproductive.

XX

OS Homo sapiens.

XX

FN WO200200677-A1.

XX

PD 03-JAN-2002.

XX

PE 07-JUN-2001; 2001WO-US18569.

XX

PR 07-JUN-2000; 2000US-209467P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Birse CE, Rosen CA;

XX

DR WPI; 2002-147878/19.

XX

DR N-PSDB; AB054588.

XX

PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,

PT useful in the prevention, treatment and diagnosis of cancer (e.g.

PT ovarian cancer), immune disorders, cardiovascular disorders and

PT neurological diseases -

XX

PS Claim 11; SEQ ID NO 2643; 2922pp; English.

XX

CC The invention relates to 2175 novel human ovarian antigens (ABP41054-ABP43228) and to cDNAs encoding them (AB054131-AB056303), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various ovary and/or breast-related disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present

CC sequence represents a human ovarian antigen of the invention.

CC Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic format directly from WIPO CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 460 AA;

Query Match 85.1%; Score 88.5; DB 23; Length 460;

Best Local Similarity 94.7%; Pred. No. 8e-06; Indels 1; Gaps 1;

Matches 18; Conservative 0; Mismatches 0;

QY 1 AALEAKICHQIEEYFGDF 19

Db 63 AALEAKICHQI-EYFGDF 80

RESULT 7

AA43394

ID AA43394 standard; peptide; 21 AA.

XX

AC AA43394;

XX

DT 12-MAY-1994 (first entry)

XX

DE La/SSB epitope 17.

XX

KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;

KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;

KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';

KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

XX

OS Homo sapiens.

XX

FN WO9321223-A.

XX

PD 28-OCT-1993.

XX

PF 13-APR-1993; 93WO-US03484.

XX

PR 13-APR-1992; 92US-0867819.

XX

PA (OKLA) UNIV OKLAHOMA STATE.

XX

PI Harley JB;

XX

DR WPI; 1993-351658/44.

XX

PT New linear epitope(s) for human auto-antibodies - from the

PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used

PT for diagnosing and treating auto-immune disorders e.g. systemic

PT lupus erythematosus

XX

PS Claim 1; Page 30; 43pp; English.

XX

CC The sequences given in AA43391-562 are linear epitopes which are

CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,

CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'

CC polypeptide. These antigens are common in systemic lupus

CC erythematosus (SLE) and closely related disorders. The Ro/SSA family

CC of proteins has been shown to have several molecular forms which are

CC defined by the molecular weight of the antigen identified. The major

CC form has a molecular weight of 60 kD and two additional forms have

CC molecular weights of 52 and 54 kD. La/SSB is also a member of this

CC group of autoantibodies and binds small RNAs with a polyuridine

CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin

CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric

CC phosphoprotein which associates with RNA polymerase III transcripts.

CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and

CC U5 RNA. Anti-Sm antibodies may be directed against one or a

CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),

CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be

CC used for preventing, treating or screening autoimmune disorders,

CC especially SLE or Sjogrens syndrome (SS). They bind to a human

Matches	7;	Conservative	7;	Mismatches	2;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	----	--------	----	------	----

DE Lapine gram-negative lipopolysaccharide binding protein.
 XX Gram negative bacteria; LPS; high density lipoprotein;
 KW rabbit; glycoprotein; gastrointestinal tract; endotoxaemia.
 XX
 OS Oryctolagus cuniculus.

XX Key Location/Qualifiers
 FH Protein 27..482
 FT /label- LBP

XX US245013-A.

XX 14-SEP-1993.

XX 30-APR-1985; 85US-0728833.

XX 30-APR-1985; 85US-0728833.

PR 30-DEC-1986; 86US-0006710.

PR 16-JUN-1989; 89US-0367454.

XX (ULEV/) ULEVITCH R.

XX Tobias P, Ulevitch R;

XX WPI; 1993-302704/38.

DR N-PSDB; AAQ47904.

XX New glyco:protein which binds Gram negative lipo:polysaccharide -

PT isolated from acute phase sera, useful for determ. of endotoxin

PT in body fluid

XX Claim 1; Fig 17; 41pp; English.

XX A rabbit liver cDNA library was prepared using Poly-A+ RNA isolated
 CC from animals 24 hours after silver nitrate induction. The library
 CC was transferred to nitrocellulose and screened with a probe based
 CC on the N-terminal sequence of rabbit LBP and a EcoRV-cut human LBP
 CC probe. The coding sequence obtained from positive clones encodes a
 CC glycoprotein which binds to a lipopolysaccharide secreted by a
 CC gram negative bacterium. The glycoprotein retards in vitro binding
 CC of LPS to high density lipoprotein present in normal serum.

XX Sequence 482 AA;

Query Match 45.2%; Score 47; DB 14; Length 482;
 Best Local Similarity 81.8%; Pred. NO. 36;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 LEAKICHQIEE 13

DB 194 LESKICRQIEE 204

RESULT 11

AAAR53461.

ID AAR53461 standard; Protein; 482 AA.

XX AAR53461;

XX 09-DEC-1994 (first entry)

DE Lapine polysaccharide binding protein.

XX LBP; lapine lipopolysaccharide binding protein; detection;
 KW gram negative bacterial infection; ss.

OS Oryctolagus cuniculus.

XX Key Location/Qualifiers

FT Peptide 1..26

FT /label- signal peptide

FT Protein 27..451

FT /label- mature protein

XX US310879-A.

XX 10-MAY-1994.

XX 30-APR-1985; 85US-0728833.

XX 30-APR-1985; 85US-0728833.

PR 30-DEC-1986; 86US-0006710.

PR 16-JUN-1989; 89US-0367454.

PR 02-MAR-1992; 92US-0847562.

XX (SCRI) SCRIPPS RES INST.

XX Tobias PS, Ulevitch RJ;

DR WPI; 1994-150499/18.

DR N-PSDB; AAQ74305.

XX Antibodies immunoreactive with lapine lipopolysaccharide binding
 PT protein - useful for determ. of presence and amt. of human and
 PT rabbit LBP.

XX Disclosure; Fig 17; 44pp; English.

XX AAR53461 shows a lapine lipopolysaccharide binding protein (LBP).
 CC The lapine LBP was used to raise antibodies against it to further
 CC identify other LBPs.

CC LBP is a glycoprotein that binds lipopolysaccharide that can be
 CC present in the bloodstream of an individual susceptible to infection
 CC by LPS-secreting gram negative bacteria.

XX Sequence 482 AA;

Query Match 45.2%; Score 47; DB 15; Length 482;
 Best Local Similarity 81.8%; Pred. No. 36;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 LEAKICHQIEE 13

DB 194 LESKICRQIEE 204

RESULT 12

AAW40813

ID AAW40813 standard; Protein; 482 AA.

XX AAW40813;

XX 01-APR-1998 (first entry)

DE Rabbit lipopolysaccharide binding protein.

XX Lipopolysaccharide binding protein; LBP; antibody; acute phase; LPS;
 KW gram-negative bacterial infection; tissue injuring infection; rabbit;
 KW surgical trauma; tissue infarction; idiopathic inflammatory state.

XX Oryctolagus cuniculus.

XX US5698680-A.

XX 16-DEC-1997.

XX 10-MAY-1994; 94US-0240328.

XX 02-MAR-1992; 92US-0847562.

PR 30-APR-1985; 85US-0728833.

PR 30-DEC-1986; 86US-0006710.

PR 16-JUN-1989; 89US-0367454.

PR 10-MAY-1994; 94US-0240328.

XX (SCRI) SCRIPPS RES INST.

XX PT Tobias PS, Ulevitch RJ;
 XX DR WPI; 1998-051547/05.
 XX DR N-PSDB; AAT99249.
 XX PT Anti-lipo-polysaccharide binding protein, LBP, antibody - useful in
 PT immunoassay for LBP, which is present in acute phase serum, but
 PT absent from normal serum
 XX PS Disclosure; Fig 17; 42pp; English.
 XX CC This sequence represents the rabbit lipopolysaccharide binding protein
 CC (LBP). This protein is used to produce the anti-LBP antibodies (Ab) of
 CC the invention. The Ab can be used in immunoassays for LBP, which is
 CC present in serum from the acute phase of gram-negative bacterial
 CC infection, but absent from normal serum. Injuries that lead to an acute
 CC phase response include tissue injuring infection, surgical or other
 CC trauma, drug related effects, burns, tissue infarction and idiopathic
 CC inflammatory states. The LBP protein can be used in an assay for
 CC detecting the presence of a pharmacologically active lipopolysaccharide
 CC (LPS) which is released from the cell wall during gram-negative bacterial
 CC infection.
 XX CC
 XX SQ Sequence 482 AA;
 Query Match 45.2%; Score 47; DB 19; Length 482;
 Best Local Similarity 81.8%; Pred. No. 36;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 3 LEAKICHOIEE 13
 II:III IIII
 DB 194 LESKICRQIEE 204
 RESULT 13
 AAW78355
 ID AAW78355 standard; Protein; 482 AA.
 XX AC AAW78355;
 XX DT 11-MAY-1999 (first entry)
 XX DE Rabbit lipopolysaccharide binding protein.
 KW Lipopolysaccharide binding protein; LPS; LBP; rabbit; serum; assay;
 KW high density lipoprotein; Gram-negative bacteria; infection.
 XX OS Oryctolagus cuniculus.
 XX FH Key Location/Qualifiers
 FT Peptide 1..26
 FT Protein /note= "signal peptide"
 FT /note= "mature LPS-binding protein"
 XX US5871937-A.
 XX PN 16-FEB-1999.
 XX PD 15-DEC-1997; 97US-0990849.
 XX PF 02-MAR-1992; 92US-0847562.
 XX PR 30-APR-1985; 85US-0728833.
 XX PR 30-DEC-1986; 86US-0006710.
 XX PR 16-JUN-1989; 89US-0367454.
 XX PR 10-MAY-1994; 94US-0240328.
 XX PR 15-DEC-1997; 97US-0990849.
 XX (SCRI) SCRIPPS RES INST.
 XX Tobias PS, Ulevitch RJ;
 XX

DR WPI; 1999-166631/14.
 DR N-PSDB; AAX16965.
 XX Detecting lipo-poly-saccharide binding protein in mammals -
 PT comprises contacting sample with binding ligand, and detecting
 PT resulting complex
 XX PS Claim 9; Fig 17; 43pp; English.
 XX CC This sequence represents a lipopolysaccharide (LPS)-binding protein (LBP)
 CC from rabbits. LBP is an acute phase protein that binds to LPS and retards
 CC binding of LPS to high density lipoprotein in animal serum. The protein
 CC can be used in assays to detect the presence of LPS from bacteria,
 CC especially Gram-negative bacteria, as a result of bacterial infection.
 XX SQ Sequence 482 AA;
 Query Match 45.2%; Score 47; DB 20; Length 482;
 Best Local Similarity 81.8%; Pred. No. 36;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 3 LEAKICHOIEE 13
 II:III IIII
 DB 194 LESKICRQIEE 204
 RESULT 14
 AAB51905
 ID AAB51905 standard; Protein; 132 AA.
 XX AC AAB51905;
 XX DT 16-FEB-2001 (first entry)
 XX DE Human secreted protein sequence encoded by gene 23 SEQ ID NO:138.
 KW Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
 KW antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective;
 KW neurotropic; neuroprotective; antibacterial; virucide; fungicide;
 KW opthalmological; vulnery; autoimmune disease; rheumatoid arthritis;
 KW hyperproliferative disorders; cancer; cardiovascular disorder;
 KW cardiac arrest; cerebrovascular disorder; nervous system disorder;
 KW Alzheimer's disease; ocular disorder; wound healing; skin aging.
 XX OS Homo sapiens.
 XX PN WO2000061626-A1.
 XX PD 19-OCT-2000.
 XX PF 06-APR-2000; 2000WO-US09066.
 XX PR 09-APR-1999; 99US-0128698.
 XX PR 20-JAN-2000; 2000US-0176926.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX PA (ROSE/) ROSEN C A.
 XX PI Rosen CA, Ruben SM, Komatsoulis G;
 XX WPI; 2000-619227/59.
 DR N-PSDB; AAC93501.
 XX New nucleic acid molecules encoding 49 human secreted proteins for
 PT diagnosing, preventing or ameliorating medical conditions and used for
 PT food additives or preservatives -
 XX Disclosure; Page 63; 516pp; English.
 XX Polynucleotide sequences AAC93479 - AAC93527 represent cDNA encoding
 CC human secreted proteins AAB51827 - AAB51875. Sequences AAB51876 -
 CC AAB51927 represent alternative polypeptides encoded by the genes, and
 CC amino acid sequences with which they share homology. The genes and

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:34:48 ; Search time 21.1348 Seconds
(without alignments)
185.234 Million cell updates/sec

Title: US-09-836-073-16
Perfect score: 104
Sequence: 1 AALEAKICHQIEEYFGDF 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 9297

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriaph.*
17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	29.8	20	Q95MK4	Q95mk4 cheirogaleu
2	28	26.9	11	Q9UC46	Q9uc46 homo sapien
3	28	26.9	20	Q9R4L7	Q9r4l7 leuconostoc
4	27	26.0	15	Q9TR89	Q9tr89 bos taurus
5	27	26.0	18	Q16028	Q16028 homo sapien
6	27	26.0	20	Q95MJ9	Q95mj9 pan troglod
7	27	26.0	22	Q9PXA8	Q9pxa8 human immun
8	27	26.0	24	Q28430	Q28430 gorilla gor
9	27	26.0	24	Q28781	Q28781 pongo pygma
10	27	26.0	24	Q38270	Q38270 bacterioph
11	27	26.0	24	Q9PXA7	Q9pxa7 human immun
12	26	25.0	8	Q9BYY5	Q9byy5 homo sapien
13	26	25.0	8	Q9BFA0	Q9bfa0 macaca mula
14	26	25.0	8	Q9BF99	Q9bf99 hylobates c
15	26	25.0	15	Q9S8M9	Q9s8m9 raphanus sa
16	26	25.0	20	Q95MK0	Q95mk0 pongo pygma

17	26	25.0	24	2	Q05616	Q05616 staphylococ
18	26	25.0	25	8	Q9TGB8	Q9tgb8 alnus crisp
19	26	25.0	25	8	Q9TGB7	Q9tgb7 alnus glut
20	26	25.0	25	8	Q9TGB6	Q9tgb6 alnus marit
21	26	25.0	25	8	Q9TGB5	Q9tgb5 betula alle
22	26	25.0	25	8	Q9TGB4	Q9tgb4 betula glan
23	26	25.0	25	8	Q9TGB3	Q9tgb3 betula papy
24	26	25.0	25	8	Q9TGB2	Q9tgb2 betula verr
25	26	25.0	25	8	Q9TGB1	Q9tgb1 betula pube
26	26	25.0	25	8	Q9TGB0	Q9tgb0 corylus ave
27	26	25.0	25	8	Q9TGA9	Q9tga9 corylus col
28	26	25.0	25	8	Q9TGA8	Q9tga8 corylus cor
29	26	25.0	25	8	Q9TGA7	Q9tga7 ostriva virg
30	26	25.0	25	8	Q9TGA6	Q9tga6 quercus rub
31	26	25.0	25	8	Q9TGB3	Q9tgb3 carpinus ca
32	26	25.0	25	13	Q90Z49	Q90z49 haplochromi
33	25	24.0	13	8	Q9THR8	Q9thr8 bryopsis sp
34	25	24.0	13	12	Q9EIV4	Q9eiv4 hepatitis b
35	25	24.0	13	12	Q9EIV3	Q9eiv3 hepatitis b
36	25	24.0	13	12	Q9EIV2	Q9eiv2 hepatitis b
37	25	24.0	13	12	Q9EIV1	Q9eiv1 hepatitis b
38	25	24.0	13	12	Q9EIV0	Q9eiv0 hepatitis b
39	25	24.0	13	12	Q9EIU9	Q9eiu9 hepatitis b
40	25	24.0	13	12	Q9EIU8	Q9eiu8 hepatitis b
41	25	24.0	13	12	Q9EIU6	Q9eiu6 hepatitis b
42	25	24.0	13	12	Q9EIU5	Q9eiu5 hepatitis b
43	25	24.0	13	12	Q9EIU4	Q9eiu4 hepatitis b
44	25	24.0	13	12	Q9EIU3	Q9eiu3 hepatitis b
45	25	24.0	13	12	Q9EIU2	Q9eiu2 hepatitis b

ALIGNMENTS

RESULT 1

Q95MK4 PRELIMINARY; PRT; 20 AA.
ID. Q95MK4
AC Q95MK4
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Growth hormone receptor (Fragment).
OS Cheirogaleus major (Lemur).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Cheirogaleidae;
OC Cheirogaleus.
OX NCBI_TaxID=47177;
RN [1]
RP SEQUENCE FROM N.A. PubMed=11371582;
RX MEDLINE=21265430; PubMed=11371582;
RA Liu J.C., Makova K.D., Adkins R.M., Gibson S., Li W.H.;
RT "Episodic Evolution of Growth Hormone in Primates and Emergence of the
RT Species Specificity of Human Growth Hormone Receptor.";
RL Mol. Biol. Evol. 18:945-953(2001).
DR EMBL; AF339072; AAK62299.1, -
KW Receptor.
FT NON_TER 1 1
SQ SEQUENCE 20 AA; 2381 MW; AAF9F94636B54BFE CRC64;
Query Match 29.8%; Score 31; DB 6; Length 20;
Best Local Similarity 40.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 LEAKICHQIEEYFG 17
||| : : : :
DB 5 LETFSCHWTDEVHVG 19

RESULT 2

Q9UC46 PRELIMINARY; PRT; 11 AA.
ID Q9UC46
AC Q9UC46;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE Neutrophil inhibitor peptide, NIP-POLYMORPHONUCLEAR neutrophil
 DE Inhibitor peptide.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96326114; PubMed=8703476;
 RA Cooper J.A.Jr., Culbreth R.R.;
 RT "Characterization of a neutrophil inhibitor peptide harvested from
 human bronchial lavage: homology to influenza A nucleoprotein.";
 RL Am. J. Respir. Cell Mol. Biol. 15:207-215(1996).
 SQ SEQUENCE 11 AA; 1262 MW; 951A1C3279C9DB45 CRC64;

 Query Match 26.9%; Score 28; DB 4; Length 11;
 Best Local Similarity 80.0%; Pred. No. 4.6e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 14 YFEGD 18
 DB 5 YFEGD 9
 ID Q9R4L7 PRELIMINARY; PRT; 20 AA.
 AC Q9R4L7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE Dextranase (EC 2.4.1.5) (Fragment).
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95306915; PubMed=7540436;
 RA Funane K., Yamada M., Shiraiwa M., Takahara H., Yamamoto N.,
 RA Ichishima E., Kobayashi M.;
 RT "Aggregated form of dextranases from Leuconostoc mesenteroides
 NRRL B-512F and its constitutive mutant.";
 RL Biosci. Biotechnol. Biochem. 59:776-780(1995).
 DR InterPro: IPR003318; GH.70.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 20 AA; 2356 MW; 701179A7AC89957F CRC64;

 Query Match 26.9%; Score 28; DB 2; Length 20;
 Best Local Similarity 44.4%; Pred. No. 8.6e+02;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

 QY 10 QIEEYFGD 18
 DB 2 QVPRVYGD 10
 ID Q9TR89 PRELIMINARY; PRT; 15 AA.
 AC Q9TR89;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE PA28 protein (Fragment).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21265430; PubMed=11371582;
 RA Liu J.C., Makova K.D., Adkins R.M., Gibson S., Li W.H.;
 RT "Episodic Evolution of Growth Hormone in Primates and Emergence of the
 RT Species Specificity of Human Growth Hormone Receptor.";
 RL Mol. Biol. Evol. 18:945-953(2001).
 DR EMBL; AF339077; AAK62304.1; -.
 KW Receptor.

RN [1]
 RP SEQUENCE.
 RX MEDLINE=95081084; PubMed=7989312;
 RA Mott J.D., Pramanik B.C., Moomaw C.R., Afendis S.J., DeMartino G.N.,
 RA Slaugther C.A.;
 RT "PA28, an activator of the 20 S proteasome, is composed of two
 RT nonidentical but homologous subunits.";
 RL J. Biol. Chem. 269:31466-31471(1994).
 SQ SEQUENCE 15 AA; 1743 MW; 61145E24192E1DE CRC64;

 Query Match 26.0%; Score 27; DB 6; Length 15;
 Best Local Similarity 33.3%; Pred. No. 9.3e+02;
 Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

 QY 7 IQHIEYFGD 18
 DB 2 LVHDEAVYGD 13
 ID Q16028 PRELIMINARY; PRT; 18 AA.
 AC Q16028;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DE OCRL-1 protein (Fragment).
 DE OCRL-1.
 GN OCRL-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93278398; PubMed=8504307;
 RA Leahy A.M., Charnas L.R., Nusebaum R.L.;
 RT "Nonsense mutations in the OCRL-1 gene in patients with the
 RT oculocerebrorenal syndrome of Lowe.";
 RL Hum. Mol. Genet. 2:461-463(1993).
 DR EMBL; S62084; AAD13933.1; -.
 FT NON_TER 1
 SQ SEQUENCE 18 AA; 2055 MW; 61356C825F7AF33D CRC64;

 Query Match 26.0%; Score 27; DB 4; Length 18;
 Best Local Similarity 36.4%; Pred. No. 1.1e+03;
 Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

 QY 1 AALEAKICHOI 11
 DB 1 SAYDPRICROL 11
 ID Q95MJ9 PRELIMINARY; PRT; 20 AA.
 AC Q95MJ9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DE Growth hormone receptor (Fragment).
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21265430; PubMed=11371582;
 RA Liu J.C., Makova K.D., Adkins R.M., Gibson S., Li W.H.;
 RT "Episodic Evolution of Growth Hormone in Primates and Emergence of the
 RT Species Specificity of Human Growth Hormone Receptor.";
 RL Mol. Biol. Evol. 18:945-953(2001).
 DR EMBL; AF339077; AAK62304.1; -.
 KW Receptor.

```

FT  NON_TER      1      1
ET  NON_TER      20     20
SQ  SEQUENCE     20 AA; 2412 MW; 3BD9F94636B54BE7 CRC64;

Query Match
Best Local Similarity 40.0%; Score 27; DB 6; Length 20;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY  8 CHQIEEYFG 17
    || : | : |
Db  10 CHTWDEVHVG 19

RESULT 7
Q9PXA8
ID  Q9PXA8      PRELIMINARY;      PRT;      22 AA.
AC  Q9PXA8;
DT  01-MAY-2000 (TrEMBLrel. 13, Created)
DT  01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT  01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE  A/ASE protein (Fragment).
OS  Human immunodeficiency virus type 1.
OC  Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX  NCBI_TaxID=11676;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=95100972; PubMed=7802674;
RA  Salberg M., Sherefa K., Zhang Z.X.;
RT  "The antigen/antibody specificity exchanger: a new peptide based tool
RT  for re-directing antibodies of other specificities to recognize the V3
RT  domain of HIV-1 gp120.";
RL  Blochem. Biophys. Res. Commun. 205:1386-1390(1994).
SQ  SEQUENCE 22 AA; 2704 MW; 36E8F13D871579CE CRC64;

Query Match
Best Local Similarity 26.0%; Score 27; DB 15; Length 22;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  12 EEYVF 16
    | : | |
Db  10 EDYVF 14

RESULT 8
Q28430
ID  Q28430      PRELIMINARY;      PRT;      24 AA.
AC  Q28430;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE  Relaxin (Fragment).
GN  RLX.
OS  Gorilla gorilla (gorilla).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX  NCBI_TaxID=9593;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=RG YK;
RA  Evans B.B.A., Fu P., Tregear G.G.W.;
RT  "Characterisation of primate relaxin genes.";
RL  J. Mol. Endocrinol. 0:0-0(1993).
CC  -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR  EMBL; 227228; CA81742.1; -.
DR  HSSP; P04090; 6RLX.
DR  InterPro; IPR004825; Ins/IGF/relax.
DR  Pfam; PF00049; Insulin; 1.
DR  PROSITE; PS00262; INSULIN; 1.
FT  NON_TER      1
SQ  SEQUENCE 24 AA; 2673 MW; 5615D4B92C1DFD24 CRC64;

Query Match
Best Local Similarity 26.0%; Score 27; DB 6; Length 24;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY  1 AALEAKICH 9
    : | | | |
Db  4 SALANKCCH 12

RESULT 10
Q38270
ID  Q38270      PRELIMINARY;      PRT;      24 AA.
AC  Q38270;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE  Bacteriophage lambda mutant truncated regulatory protein (PyrI) gene
DE  from E.coli, 3' end (Fragment).
OS  Bacteriophage lambda.
OC  Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC  Lambda-like viruses.
OX  NCBI_TaxID=10710;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  MEDLINE=86143826; PubMed=3912513;
RA  Cunin R., Jacobs A., Charlier D., Crabeel M., Herve G., Glansdorff N.,
RA  Pierard A.;
RT  "Structure-function relationship in allosteric aspartate
RT  carbamoyltransferase from Escherichia coli: I. Primary structure of a
RT  pyrI gene encoding a modified regulatory subunit.";
RL  J. Mol. Biol. 186:707-713(1985).
DR  EMBL; M28579; AAA32252.1; -.
DR  HSSP; P00478; 8ATC.
FT  NON_TER      1
SQ  SEQUENCE 24 AA; 2905 MW; FD9349DF6F8159D0 CRC64;

Query Match
Best Local Similarity 26.0%; Score 27; DB 6; Length 24;

```

```

Best Local Similarity 55.6%; Pred. No. 1.5e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY  1 AALEAKICH 9
    : | | | |
Db  4 SALANKCCH 12

RESULT 9
Q28781
ID  Q28781      PRELIMINARY;      PRT;      24 AA.
AC  Q28781;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE  Relaxin (Fragment).
GN  RLX.
OS  Pongo pygmaeus (Orangutan).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX  NCBI_TaxID=9600;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=RG YK 287;
RA  Evans B.B.A., Fu P., Tregear G.G.W.;
RT  "Characterisation of primate relaxin genes.";
RL  J. Mol. Endocrinol. 0:0-0(1993).
CC  -1- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR  EMBL; 227229; CA81743.1; -.
DR  HSSP; P04090; 6RLX.
DR  InterPro; IPR004825; Ins/IGF/relax.
DR  Pfam; PF00049; Insulin; 1.
DR  PROSITE; PS00262; INSULIN; 1.
FT  NON_TER      1
SQ  SEQUENCE 24 AA; 2673 MW; 5615D4B92C1DFD24 CRC64;

Query Match
Best Local Similarity 55.6%; Pred. No. 1.5e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY  1 AALEAKICH 9
    : | | | |
Db  4 SALANKCCH 12

RESULT 10
Q38270
ID  Q38270      PRELIMINARY;      PRT;      24 AA.
AC  Q38270;
DT  01-NOV-1996 (TrEMBLrel. 01, Created)
DT  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT  01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE  Bacteriophage lambda mutant truncated regulatory protein (PyrI) gene
DE  from E.coli, 3' end (Fragment).
OS  Bacteriophage lambda.
OC  Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC  Lambda-like viruses.
OX  NCBI_TaxID=10710;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  MEDLINE=86143826; PubMed=3912513;
RA  Cunin R., Jacobs A., Charlier D., Crabeel M., Herve G., Glansdorff N.,
RA  Pierard A.;
RT  "Structure-function relationship in allosteric aspartate
RT  carbamoyltransferase from Escherichia coli: I. Primary structure of a
RT  pyrI gene encoding a modified regulatory subunit.";
RL  J. Mol. Biol. 186:707-713(1985).
DR  EMBL; M28579; AAA32252.1; -.
DR  HSSP; P00478; 8ATC.
FT  NON_TER      1
SQ  SEQUENCE 24 AA; 2905 MW; FD9349DF6F8159D0 CRC64;

```

Query Match 26.0%; Score 27; DB 9; Length 24;
 Best Local Similarity 42.9%; Pred. No. 1.5e+03;
 Matches 6; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

QY 2 ALEAKICHQIEEY 15
 ||:|:|:|:|

Db 8 ALCKRYCE--KEFY 19

RESULT 11

Q9PXA7 PRELIMINARY; PRT; 24 AA.
 AC Q9PXA7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE A/ASE protein (Fragment).
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95100972; PubMed=7802674;
 RA Salberg M., Sheraga K., Zhang Z.X.;
 RT "The antigen/antibody specificity exchanger: a new peptide based tool
 for re-directing antibodies of other specificities to recognize the V3
 domain of HIV-1 gp120."
 RT Blochem. Biophys. Res. Commun. 205:1386-1390(1994).
 RL InterPro: IPR002006; Hepatitis_Core.
 DR Pfam: PF00906; Hepatitis_Core; 1.
 SQ SEQUENCE 24 AA; 2961 MW; 0A3B5E7F96173E96 CRC64;

Query Match 26.0%; Score 27; DB 15; Length 24;
 Best Local Similarity 80.0%; Pred. No. 1.5e+03;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYVF 16
 |:|:|

Db 10 EDYF 14

RESULT 12

Q9BY5 PRELIMINARY; PRT; 8 AA.
 AC Q9BY5;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CAMP responsive element moderator (Fragment).
 GN CREM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21082082; PubMed=11214319;
 RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
 O'Brien S.J.;
 RT "Molecular phylogenetics and the origins of placental mammals.";
 RL Nature 409:614-618(2001).
 DR EMBL: AY011664; AAG47575.1; -.
 FT NON_TER 1 1
 SQ SEQUENCE 8 AA; 1006 MW; DF02C331EEAB572A CRC64;

Query Match 25.0%; Score 26; DB 4; Length 8;
 Best Local Similarity 60.0%; Pred. No. 6.7e+05;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHOIE 12
 |:|:|

Db 4 CHKVE 8

RESULT 13

Q9BFA0 PRELIMINARY; PRT; 8 AA.
 AC Q9BFA0;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CAMP responsive element moderator (Fragment).
 GN CREM.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheidae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21082082; PubMed=11214319;
 RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
 O'Brien S.J.;
 RT "Molecular phylogenetics and the origins of placental mammals.";
 RL Nature 409:614-618(2001).
 DR EMBL: AY011661; AAG47572.1; -.
 FT NON_TER 1 1
 SQ SEQUENCE 8 AA; 1006 MW; DF02C331EEAB572A CRC64;

Query Match 25.0%; Score 26; DB 6; Length 8;
 Best Local Similarity 60.0%; Pred. No. 6.7e+05;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHOIE 12
 |:|:|

Db 4 CHKVE 8

RESULT 14

Q9BF99 PRELIMINARY; PRT; 8 AA.
 AC Q9BF99;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CAMP responsive element moderator (Fragment).
 GN CREM.
 OS Hylobates concolor (crested gibbon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
 OX NCBI_TaxID=29089;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21082082; PubMed=11214319;
 RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
 O'Brien S.J.;
 RT "Molecular phylogenetics and the origins of placental mammals.";
 RL Nature 409:614-618(2001).
 DR EMBL: AY011662; AAG47573.1; -.
 FT NON_TER 1 1
 SQ SEQUENCE 8 AA; 1006 MW; DF02C331EEAB572A CRC64;

Query Match 25.0%; Score 26; DB 6; Length 8;
 Best Local Similarity 60.0%; Pred. No. 6.7e+05;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHOIE 12
 |:|:|

Db 4 CHKVE 8

RESULT 15

Q9S8M9 PRELIMINARY; PRT; 15 AA.
 AC Q9S8M9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Ascorbate peroxidase homology (Fragment).
 OS Raphanus sativus (Radish).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Raphanus.
 OX NCBI_TaxID=3726;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=94272301; PubMed=8003938;
 RA Lopez F.; Vansuyt G.; Derancourt J.; Fourcroy P.; Casse-Delbart F.;
 RT "Identification by 2D-page analysis of salt-stress induced proteins in
 RT radish (Raphanus sativus).";
 RL Cell. Mol. Biol. 40:85-90(1994).
 SQ SEQUENCE 15 AA; 1683 MW; D38B598E5EF892B4 CRC64;

Query Match 25.0%; Score 26; DB 10; Length 15;
 Best Local Similarity 50.0%; Pred. No. 1.4e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 12 EYFEGDF 19
 |||||
 Db 5 EEAFFADY 12

Search completed: April 23, 2003, 13:47:16
 Job time : 21.1348 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:33:13 ; Search time 4.91011 Seconds

(without alignments)
160.495 Million cell updates/sec

Title: US-09-836-073-16

Perfect score: 104

Sequence: 1 AALEAKICHQIEEYFFGDF 19

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1520

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	26.0	19	1	MBB2_UROHA
2	26	25.0	16	1	MLB_SQAC
3	24	23.1	11	1	CXLI_CONMR
4	24	23.1	12	1	TIN2_HOPTI
5	24	23.1	13	1	CXLA_CONMR
6	24	23.1	20	1	COXN_THUOB
7	24	23.1	21	1	PEDB_HYDAT
8	23.5	22.6	18	1	SFAH_HELAN
9	23.5	22.6	23	1	XYCI_ACIGB
10	23	22.1	25	1	FRHB_METBA
11	22	21.2	15	1	PGTS_PELAC
12	22	21.2	21	1	BTX_ATRBI
13	22	21.2	21	1	SRD_ATREN
14	22	21.2	23	1	NUO5_SOLTU
15	22	21.2	25	1	ANDT_ANDAU
16	21	20.2	7	1	ALL2_CARMA
17	21	20.2	7	1	ALL3_CARMA
18	21	20.2	23	1	CLVB_STYCL
19	21	20.2	25	1	SMBP_RAT
20	20	19.2	9	1	FIBB_EPYPA
21	20	19.2	15	1	LECL_PSOSC
22	20	19.2	15	1	PC20_BRANA
23	20	19.2	16	1	CXA2_CONMA
24	20	19.2	20	1	ALAT_PIG
25	20	19.2	20	1	KORA_METTM
26	20	19.2	21	1	NDK_CANAL
27	20	19.2	22	1	LPI_TRIWA
28	20	19.2	22	1	LP2_TRIWA
29	19	18.3	8	1	ALI7_CARMA
30	19	18.3	9	1	ALI1_CARMA
31	19	18.3	10	1	GON2_CHEPR
32	19	18.3	11	1	LSK1_LEUMA
33	19	18.3	11	1	TIN4_HOPTI

34 19 18.3 12 1 NO40_SESRO
35 19 18.3 12 1 TIN3_HOPTI
36 19 18.3 13 1 NPY_LYMST
37 19 18.3 13 1 NP2_LYMST
38 19 18.3 13 1 NP4_LYMST
39 19 18.3 13 1 NP5_LYMST
40 19 18.3 14 1 MAST_VESLE
41 19 18.3 19 1 FIBB_VULVU
42 19 18.3 20 1 COXF_ONCMY
43 19 18.3 20 1 CS21_STRTR
44 19 18.3 20 1 TL18_SPIOL
45 19 18.3 21 1 TL19_ARATH

O24369 sesbania ro
P82653 hoplobatrach
P80178 lymnaea sta
P80179 lymnaea sta
P80181 lymnaea sta
P80182 lymnaea sta
P01514 vespula lew
P14482 vulpes vulp
P80329 oncorhynchu
P81621 streptococc
P82336 spinacia ol
P82658 arabidopsis

ALIGNMENTS

RESULT 1
HBB2_UROHA
ID HBB2_UROHA STANDARD; PRT; 19 AA.
AC P18992;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hemoglobin beta-2 chain (Fragment).
OS Uromastix hardwickii (Indian spiny-tailed lizard).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Uromastycinae;
OC Uromastix.
OX NCBI_TaxID=40250;
RN [1]
RP SEQUENCE.
RX MEDLINE=84029159; PubMed=6628672;
RA Naqvi S., Zaidi Z.H., von Bahr-Lindstroem H., Carlquist M.,
RA Joernvall H.;
RT "Characterization of hemoglobin from the lizard Uromastix
hardwickii.";
RL FEBS Lett. 162:290-295(1983).
CC -!- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CC -!- VARIOUS PERIPHERAL TISSUES.
CC -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -!- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -!- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR; A05305; A05305.
DR InterPro; IPR000971; Globin.
DR PROSITE; PS01033; GLOBIN; PARTIAL.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT NON_TER 1
FT NON_TER 19
SQ SEQUENCE 19 AA; 1914 MW; C40AD8EA30019057 CRC64;

Query Match 26.0%; Score 27; DB 1; Length 19;
Best Local Similarity 80.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 15 YFGDF 19
Db 1 FFGDF 5

RESULT 2
MLB_SQAC
ID MLB_SQAC STANDARD; PRT; 16 AA.
AC P01207;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanotropin beta.

OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]

RP SEQUENCE.
RX MEDLINE=75127390; PubMed=4375978;
RA Bennett H.P.J., Lowry P.J., McMartin C., Scott A.P.;
RT "Structural studies of alpha-melanocyte-stimulating hormone and a
RT novel beta-melanocyte-stimulating hormone from the neurointermediate
RL lobe of the pituitary of the dogfish *Squalus acanthias*.";
RL Biochem. J. 141:439-444(1974).
CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
CC PIR: A01471; MTFBBS.
KW Hormone.
-SQ SEQUENCE 16 AA; 1939 MW; 993AF43C8A640A0E CRC64;

Query Match 25.0%; Score 26; DB 1; Length 16;
Best Local Similarity 50.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 12 EYFEGDF 19
: : : : :
Db 3 DDYKFGHF 10

RESULT 3
CX1L_CONMR STANDARD; PRT; 11 AA.
ID CX1L_CONMR
AC P58807;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda-conotoxin CMVIA.
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE-Venom;
RX MEDLINE=20564325; PubMed=10988292;
RA Balaji R.A., Ohtake A., Sato K., Gopalakrishnakone P., Kini R.M.,
RA Seow K.T., Bay B.-H.;
RT "Lambda-conotoxins, a new family of conotoxins with unique disulfide
RT pattern and protein folding. Isolation and characterization from the
RT venom of Conus marmoreus.";
RL J. Biol. Chem. 275:39516-39522(2000).
CC -1- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- MASS SPECTROMETRY: MW=1237.93; MW_ERR=0.21; METHOD-Electrospray.
CC -1- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
KW Neurotoxin; Toxin; Hydroxylation.
FT DISULFID 2 11
FT MOD_RES 3 8
FT SEQUENCE 11 AA; 1226 MW; 277AAC60B7232B58 CRC64;

Query Match 23.1%; Score 24; DB 1; Length 11;
Best Local Similarity 75.0%; Pred. No. 5.4e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 KICH 9
: : :
Db 6 KLCH 9

RESULT 4
TIN2_HOPTI STANDARD; PRT; 12 AA.
ID TIN2_HOPTI
AC P82652;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tigerin-2.
OS Hoplobatrachus tigerinus (Indian bull frog) (Rana tigerina).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae;
OC Hoplobatrachus.
OX NCBI_TaxID=103373;
RN [1]
RP SEQUENCE, FUNCTION, MASS SPECTROMETRY, AND DISULFIDE BONDS.
RC TISSUE-Skin;
RX PubMed=11031261;
RA Purna Sai K., Jagannadham M.V., Vairamani M., Raju N.P.,
RA Devi A.S., Nagaraj R., Sitarum N.;
RT "Tigerinins: novel antimicrobial peptides from the Indian frog *Rana*
RT *tigerina*.";
RL J. Biol. Chem. 276:2701-2707(2001).
CC -1- FUNCTION: ANTIBACTERIAL ACTIVITY AGAINST *B.SUBTILIS*, *E.COLI*,
CC *S.AUREUS*, *M.LUTEUS*, *P.PUTIDA* AND *S.CEREVISIAE*.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: SKIN.
CC -1- MASS SPECTROMETRY: MW=1368; METHOD-MALDI.
KW Amphibian skin; Antibiotic; Amidation.
FT DISULFID 3 11
FT MOD_RES 12 12
FT SEQUENCE 12 AA; 1369 MW; C2627D8EFD37605D CRC64;

Query Match 23.1%; Score 24; DB 1; Length 12;
Best Local Similarity 50.0%; Pred. No. 5.9e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 ALEAKICH 9
: : : : :
Db 5 AIPLDICH 12

RESULT 5
CX14_CONMR STANDARD; PRT; 13 AA.
ID CX14_CONMR
AC P58810;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lambda/chi-conotoxin MrIB (Chi-MrIB).
OS Conus marmoreus (Marble cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=42752;
RN [1]
RP SEQUENCE, SYNTHESIS, AND STRUCTURE BY NMR.
RC TISSUE-Venom;
RX MEDLINE=21419681; PubMed=11528421;
RA Sharpe I.A., Gehrmann J., Loughnan M.L., Thomas L., Adams D.A.,
RA Atkins A., Palant E., Craik D.J., Adams D.J., Alewood P.F.,
RA Lewis R.J.;
RT "Two new classes of conopeptides inhibit the alpha1-adrenoceptor and
RT noradrenaline transporter.";
RL Nat. Neurosci. 4:902-907(2001).
CC -1- FUNCTION: Inhibits the neuronal noradrenaline transporter.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -1- PIR: Exists in two forms, due to cis-trans isomerization at His-
CC 11-Hyp-12.
CC -1- MASS SPECTROMETRY: MW=1393.52; METHOD-Electrospray.
CC -1- SIMILARITY: BELONGS TO THE CHI/LAMBDA-CONOTOXIN FAMILY.
DR PDB; IIEO; 03-APR-02.
KW Neurotoxin; Toxin; Hydroxylation; 3D-structure.
FT DISULFID 4 13
FT DISULFID 5 10
FT MOD_RES 12 12
FT SEQUENCE 13 AA; 1382 MW; 277AAC376EAD2B58 CRC64;

Query Match 23.1%; Score 24; DB 1; Length 13;
Best Local Similarity 75.0%; Pred. No. 6.4e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 KICH 9
: : :
Db 6 KICH 9

```

Db      8 KLCH 11
RESULT 6
COXN_THUOB
ID COXN_THUOB STANDARD; PRT; 20 AA.
AC P80980;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Cytochrome c oxidase polypeptide Viib-heart (EC 1.9.3.1) (Fragment).
OS Thunnus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=97454291; PubMed=9310366;
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
RT liver";
RL Eur. J. Biochem. 248:99-103(1997).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) -> 4 ferrocycytochrome
CC c + 2 H(2)O.
KW Oxidoreductase; Mitochondrion.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2303 MW; 0A33BD3400655AA6 CRC64;

Query Match 23.1%; Score 24; DB 1; Length 20;
Best Local Similarity 40.0%; Pred. No. 9.8e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 HOIEEYFGD 18
II : : : :
Db 5 HODFHLFYGD 14

RESULT 7
PEDB_HYDAT
ID PEDB_HYDAT STANDARD; PRT; 21 AA.
AC P80577;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Pedibin.
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroidea; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE.
RX MEDLINE=96232307; PubMed=8674432;
RA Hoffmeister S.A.H.;
RT "Isolation and characterization of two new morphogenetically active
RT peptides from Hydra vulgaris.";
RL Development 122:1941-1948(1996).
CC -!- FUNCTION: MORPHOGENETICALLY ACTIVE PEPTIDE. ACTIVE IN FOOT
CC DEVELOPMENT.
KW Morphogen.
SQ SEQUENCE 21 AA; 2322 MW; 801DA27919A917FD CRC64;

Query Match 23.1%; Score 24; DB 1; Length 21;
Best Local Similarity 42.9%; Pred. No. 1e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICHQIEE 13

```

```

Db      5 VSHELEE 11
RESULT 8
SFAH_HELAN
ID SFAH_HELAN STANDARD; PRT; 18 AA.
AC P81098;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Seed fatty acyl-ester hydrolase (EC 3.1.1.1) (Fragment).
OS Helianthus annuus (Common sunflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC Heliantheae; Helianthus.
OX NCBI_TaxID=4232;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. RUSTICA / VAR. EUROFLOR; TISSUE=Seed;
RA Beisson F., Gardies A.-M., Teissere M., Ferte N., Noat G.;
RT "An esterase neosynthesized in post-germinated sunflower seeds is
RT related to a new family of lipolytic enzymes.";
RL Plant Physiol. Biochem. 35:761-765(1997).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=95210327; PubMed=7696323;
RA Teissere M., Borel M., Caillol B., Nari J., Gardies A.-M., Noat G.;
RT "Purification and characterization of a fatty acyl-ester hydrolase
RT from post-germinated sunflower seeds.";
RL Biochim. Biophys. Acta 1255:105-112(1995).
CC -!- FUNCTION: IMPLICATED IN THE BREAKDOWN OF OIL BODY-STORED LIPIDS
CC DURING POST-GERMINATION.
CC -!- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O -> an alcohol + a
CC carboxylic anion.
CC -!- TISSUE SPECIFICITY: SEED.
CC -!- DEVELOPMENTAL STAGE: POST-GERMINATION.
CC -!- SIMILARITY: BELONGS TO THE "GDSL" FAMILY OF LIPOLYTIC ENZYMES.
DR InterPro: IPR001087; Lipase_GDSL.
DR PROSITE: PS01098; LIPASE_GDSL_SER; PARTIAL.
KW Hydrolase; Lipid degradation; Glycoprotein.
FT ACT_SITE 13 13 BY SIMILARITY.
FT NON_TER 18 18
SQ SEQUENCE 18 AA; 2057 MW; 3E9A80EB6548E862 CRC64;

Query Match 22.6%; Score 23.5; DB 1; Length 18;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 5; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 10 QIEEY-FGD 18
II : : : :
Db 3 QVPEYFPGD 12

RESULT 9
XYCL_ACIGB
ID XYCL_ACIGB STANDARD; PRT; 23 AA.
AC P46365;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Benzaldehyde dehydrogenase [NAD+] I (EC 1.2.1.28) (Fragment).
OS Acinetobacter dehydrogenase 11.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=106649;
RN [1]
RP SEQUENCE.
RC STRAIN=ATCC 11171 / NCIB 8250 / CIP 63.46 / B94;
RX MEDLINE=91113163; PubMed=1989592;
RA Chalmers R.M., Keen J.N., Fewson C.A.;
RT "Comparison of benzyl alcohol dehydrogenases and benzaldehyde

```



```

FT DISULFID 1 15 .BY SIMILARITY.
FT DISULFID 3 11 BY SIMILARITY.
SQ SEQUENCE 21 AA; 2511 MW; 83A5DFB81D036AE2 CRC64;

Query Match 21.2%; Score 22; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQ 10
DB 15 CHQ 17

RESULT 13
SRD_ATREN
ID SRD_ATREN STANDARD; PRT; 21 AA.
AC P13211;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sarafotoxin-D (SDF) (SRFX-D).
OS Atractaspis engadensis (Israeli burrowing asp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Atractaspididae; Atractaspis.
OX NCBI_TaxID=8600;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=90033283; PubMed=2509240;
RA Bdoian A., Wollberg Z., Fleming G.;
RT "SRFX-D, a new native peptide of the endothelin/sarafotoxin family.";
RL FEBS Lett. 256:1-3(1989).
CC -!- FUNCTION: VASOCONSTRICTOR ACTIVITY. THESE TOXINS CAUSE CARDIAC
ARREST PROBABLY AS A RESULT OF CORONARY VASOSPASM.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: LD(50) IS 0.35 MG/KG BY INTRAVENOUS INJECTION.
CC -!- SIMILARITY: BELONGS TO THE ENDOTHELIN/SARAFOTOXIN FAMILY.
DR PIR; S06145; S06145.
DR HSP; P13208; 1SRB.
DR InterPro; IPR001928; Endothln_tox.
DR Pfam; PF00322; endothelin; 1.
DR PRINTS; PR00365; ENDOTHELIN.
DR ProDom; PD004740; Sara/bib_toxin; 1.
DR SMART; SM00272; END; 1.
DR PROSITE; PS00270; ENDOTHELIN; 1.
KW Vasoconstrictor; Toxin.
FT DISULFID 1 15 BY SIMILARITY.
FT DISULFID 3 11 BY SIMILARITY.
SQ SEQUENCE 21 AA; 2596 MW; 83A8A04D1D536AE2 CRC64;

Query Match 21.2%; Score 22; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQ 10
DB 15 CHQ 17

RESULT 14
NUO5_SOLTU
ID NUO5_SOLTU STANDARD; PRT; 23 AA.
AC P80262;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE NADH-ubiquinone oxidoreductase 29 kDa subunit (EC 1.6.5.3)
DE (EC 1.6.99.3) (Complex I-29KD) (CI-29KD) (Fragment).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

```

```

OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. Bintje; TISSUE=Tuber;
RX MEDLINE=941124587; PubMed=8294484;
RA Herz U., Schroeder W., Liddell A., Leaver C.J., Brennicke A.,
RA Grohmann L.;
RT "Purification of the NADH:ubiquinone oxidoreductase (complex I) of
the respiratory chain from the inner mitochondrial membrane of
Solanum tuberosum.";
RL J. Biol. Chem. 269:2263-2269(1994).
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
TO BE UBIQUINONE.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
CC -!- SUBCELLULAR LOCATION: MATRIX SIDE OF THE MITOCHONDRIAL INNER
MEMBRANE.
DR PIR; C49732; C49732.
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2653 MW; 4B41B8FF83412F58 CRC64;

Query Match 21.2%; Score 22; DB 1; Length 23;
Best Local Similarity 33.3%; Pred. No. 2.3e+03;
Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 12 EBYFG 17
DB 17 KDIFYG 22

RESULT 15
ANDT_ANDAU
ID ANDT_ANDAU STANDARD; PRT; 25 AA.
AC P56684; P81616;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Androctonin.
OS Androctonus australis hector (Sahara scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Androctonus.
OX NCBI_TaxID=70175;
RN [1]
RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
RC TISSUE=Hemolymph;
RX MEDLINE=97094646; PubMed=8939880;
RA Ehret-Sabatier L., Loew D., Goyffon M., Fehlbauer P., Hoffmann J.A.,
RA van Dorsselaer A., Bulet P.;
RT "Characterization of novel cysteine-rich antimicrobial peptides from
scorpion blood.";
RL J. Biol. Chem. 271:29537-29544(1996).
RN [2]
RP SYNTHESIS OF D-AMINO ACID ENANTIOMER, AND CHARACTERIZATION.
RX MEDLINE=20115101; PubMed=10642525;
RA Hetru C., Letellier L., Oren Z., Hoffmann J.A., Shai Y.;
RT "Androctonin, a novel antimicrobial peptide from scorpion Androctonus
australis: solution structure and molecular dynamics simulations in
the presence of a lipid monolayer.";
RL J. Biomol. Struct. Dyn. 17:367-380(1999).
CC -!- FUNCTION: ACTIVE AGAINST BOTH BACTERIA (GRAM-POSITIVE AND GRAM-
NEGATIVE) AND FILAMENTOUS FUNGI. ACTS ON THE MEMBRANE OF THE

```

CC BACTERIAL CELLS. IT DESTABILIZE A MEMBRANE BY MODIFYING ITS
 CC PROPERTIES.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- MASS SPECTROMETRY: MW-3076.7; METHOD-Electrospray.
 DR PDB; IC26; 12-JAN-00.
 KW Antibiotic; Fungicide; 3D-structure.
 FT DISULFID 4 20
 FT DISULFID 10 16
 SQ SEQUENCE 25 AA; 3081 MW; D4183D6EA5AA50AA CRC64;
 Query Match 21.2%; Score 22; DB 1; Length 25;
 Best Local Similarity 50.0%; Pred. No. 2.5e+03;
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 7 ICHQIE 12
 Db 3 VCRQIK 8

Search completed: April 23, 2003, 13:43:51
 Job time : 5.01011 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:36:43 ; Search time 9.92697 Seconds
(without alignments)
183.999 Million cell updates/sec

Title: US-09-836-073-16
Perfect score: 104
Sequence: 1 AALEAKICHOIEEYFGDF 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 4984

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*

- 1: PIR1:*
- 2: PIR2:*
- 3: PIR3:*
- 4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	29.8	25	2 A60286	heat-stable serine
2	29	27.9	21	2 A60225	pyruvate dehydroge
3	29	27.9	25	2 A49038	Ig lambda chain V
4	28	26.9	18	2 PH1368	Ig heavy chain DJ
5	27	26.0	14	2 S57638	T cell receptor V-
6	27	26.0	19	2 A05305	hemoglobin beta-2
7	27	26.0	21	2 S07968	T-cell receptor be
8	27	26.0	24	2 S42780	relaxin - oranguta
9	26	25.0	16	1 MTDF5S	melanotropin beta
10	26	25.0	16	2 E53284	T-cell receptor be
11	26	25.0	20	2 H22585	R-phycoerythrin ga
12	26	25.0	22	2 S46530	actin-related prot
13	26	25.0	21	2 I73733	gene N-ras protein
14	26	25.0	23	2 PH1725	Ig heavy chain V r
15	26	25.0	25	2 C57001	endo-1,4-beta-xyla
16	25	24.0	15	2 PH1366	Ig heavy chain DJ
17	25	24.0	17	2 S57514	T cell receptor be
18	25	24.0	17	2 A61211	anantin - Streptom
19	25	24.0	18	2 PQ0072	T-cell receptor be
20	25	24.0	19	2 PQ0244	Ig heavy chain CDR
21	25	24.0	20	2 PN0133	pepsin (EC 3.4.23
22	25	24.0	21	2 S78574	protein kinase C i
23	24	23.1	12	2 S7570	T cell receptor V-
24	24	23.1	12	2 PH1454	T-cell receptor al
25	24	23.1	13	2 PH0787	T-cell receptor al
26	24	23.1	14	2 S57569	T cell receptor V-
27	24	23.1	17	2 S57555	T cell receptor V-
28	24	23.1	20	2 S77989	cytochrome-c oxida
29	24	23.1	21	2 S78575	protein kinase C i

ALIGNMENTS

RESULT 1

A60286
heat-stable serine proteinase (EC 3.4.21.-) - Thermomonospora fusca (strain YX) (f
N:Alternate names: YX-proteinase
C:Species: Thermomonospora fusca
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Dec-1996
C:Accession: A60286
R:Kristjansson, M.M.; Kinsella, J.E.
Int. J. Pept. Protein Res. 36, 201-207, 1990
A:Title: Heat stable proteinase from Thermomonospora fusca. Characterization as a
A:Reference number: A60286; MUID:91107200; PMID:2132918
A:Accession: A60286
A:Molecule type: protein
A:Residues: 1-25 <KRI>
C:Keywords: extracellular protein; hydrolase; serine proteinase

Query Match 29.8%; Score 31; DB 2; Length 25;
Best Local Similarity 44.4%; Pred. No. 2.6e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 11 IEIYVGFDF 19
DB 7 LNPYIFGNY 15

RESULT 2

A60225
pyruvate dehydrogenase (lipoamide) (EC 1.2.4.1) alpha chain - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 21-Oct-1992 #sequence_revision 21-Oct-1992 #text_change 15-Oct-1999
C:Accession: A60225
R:Lawson, R.; Aitken, A.; Yeaman, S.J.
Biochem. Soc. Trans. 11, 298-299, 1983
A:Title: Primary sequence of the N-terminal region of the alpha-subunit of pyruvate
A:Reference number: A60225
A:Accession: A60225
A:Molecule type: protein
A:Status: preliminary
A:Residues: 1-21 <LAW>
C:Superfamily: pyruvate dehydrogenase (lipoamide) alpha chain; thiamin pyrophosphat

Query Match 27.9%; Score 29; DB 2; Length 21;
Best Local Similarity 46.7%; Pred. No. 4.6e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 1 AALEAKIC--HQIEE 13
DB 5 ATFEIKKCDLHLEE 19

RESULT 3

A49038
Ig lambda chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: A49038
R:Weiss, U.; Zoebelstein, R.; Rajewsky, K.
Eur. J. Immunol. 22, 511-517, 1992
A:Title: Accumulation of somatic mutants in the B cell compartment after primary immunization
A:Reference number: A49038; MUID:92164733; PMID:1537385
A:Accession: A49038
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-25 <WEI>
A:Cross-references: GB:S85732; NID:9246314; PIDN:AAB21559.1; PID:9246315
A:Experimental source: spleen
A:Note: sequence extracted from NCBI backbone (NCBIN:85732, NCBI:P:85757)
C:Keywords: heterotetramer; immunoglobulin

Query Match 27.9%; Score 29; DB 2; Length 25;
Best Local Similarity 57.1%; Pred. No. 5.4e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 11 IEYFYG 17
: : : : :
Db 10 MERYYG 16

RESULT 4
PHI368
Ig heavy chain DJ region (clone C111-112) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PHI368
R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma
A:Reference number: PHI302; MUID:93094761; PMID:1460419
A:Accession: PHI368
A:Molecule type: DNA
A:Residues: 1-18 <WAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 26.9%; Score 28; DB 2; Length 18;
Best Local Similarity 80.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 13 EYFYG 17
: : : : :
Db 6 EYFG 10

RESULT 5
S57638
T cell receptor V-J junctional alpha chain region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57638
R:Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argae, V.P.
submitted to the EMBL Data Library, June 1995
A:Description: T cell receptor repertoire for a viral epitope in humans is diversified
A:Reference number: S57494
A:Accession: S57638
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-14 <BUR>
A:Cross-references: EMBL:Z49964; NID:9886676; PIDN:CAA90238.1; PID:9886677
C:Keywords: T-cell receptor

Query Match 26.0%; Score 27; DB 2; Length 14;
Best Local Similarity 44.4%; Pred. No. 6.4e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 HOIEEYFG 17
: : : : :
Db 6 HUGNQYFG 14

RESULT 6
A05305
hemoglobin beta-2 chain - Indian spiny-tailed lizard (fragment)
C:Species: Uromastix hardwickii (Indian spiny-tailed lizard)
C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 01-Aug-1997
C:Accession: A05305
R:Nagvi, S.; Zaidi, Z.H.; von Bahr-Lindstrom, H.; Carlquist, M.; Jorntvall, H.
FEBS Lett. 162, 290-295, 1983
A:Reference number: A91314; MUID:84029159; PMID:6628672
A:Accession: A05305
A:Molecule type: protein
A:Residues: 1-19 <NAQ>
C:Superfamily: globin; globin homology
C:Keywords: erythrocyte; oxygen carrier

Query Match 26.0%; Score 27; DB 2; Length 19;
Best Local Similarity 80.0%; Pred. No. 8.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 15 YFGDF 19
: : : : :
Db 1 FFGDF 5

RESULT 7
S07968
T-cell receptor beta chain V-J region (clone 18 BM 142) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 23-Jul-1999
C:Accession: S07968
R:Bill, J.; Vague, J.; Appel, V.B.; White, J.; Horn, G.; Erlich, H.A.; Palmer, E.
J. Exp. Med. 169, 115-133, 1989
A:Title: Molecular genetic analysis of 178 I-A(bm12)-reactive T cells.
A:Reference number: S05590; MUID:89080476; PMID:2783331
A:Accession: S07968
A:Molecule type: mRNA
A:Residues: 1-21 <BIL>
A:Cross-references: EMBL:X14937; NID:954888; PIDN:CAA33063.1; PID:9773253
A:Note: this sequence was determined from the differentiated gene
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor
F:9-21/Domain: J region (J-beta 2.2) <JRE>

Query Match 26.0%; Score 27; DB 2; Length 21;
Best Local Similarity 57.1%; Pred. No. 9.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 12 EYFYG 18
: : : : :
Db 8 EQLYFGE 14

RESULT 8
S42780
relaxin - orangutan (fragment)
C:Species: Pongo pygmaeus (orangutan)
C:Date: 08-May-1995 #sequence_revision 26-Jul-1996 #text_change 16-Jul-1999
C:Accession: S42780
R:Evans, B.B.
submitted to the EMBL Data Library, November 1993
A:Reference number: S42776
A:Accession: S42780
A:Molecule type: DNA
A:Residues: 1-24 <EVA>
A:Cross-references: EMBL:Z27229; NID:9415990; PIDN:CAA81743.1; PID:9415991
C:Genetics: rlx1
C:Superfamily: insulin

C:Keywords: disulfide bond; hormone
F:1-24/Domain: relaxin chain 1B (fragment) #status predicted <RXB1>

Query Match 26.0%; Score 27; DB 2; Length 24;
Best Local Similarity 55.6%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AALEAKICH 9
: || | ||
Db 4 SALANKCCH 12

RESULT 9
MTDPBS
melanotropin beta - spiny dogfish
C:Species: Squalus acanthias (spiny dogfish)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 23-Aug-1996
C:Accession: A01471

R:Bennett, H.P.J.; Lowry, P.J.; McMartin, C.; Scott, A.P.
Biochem. J. 141, 439-444, 1974
A:Title: Structural studies of alpha-melanocyte-stimulating hormone and a novel beta-mel
A:Reference number: A90277; MUID:75127390; PMID:4375978
A:Accession: A01471
A:Molecule type: protein
A:Residues: 1-16 <BEN>
C:Superfamily: corticotropin-lipotropin
C:Keywords: hormone

Query Match 25.0%; Score 26; DB 1; Length 16;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 12 EBYFGDF 19
: : || ||
Db 3 DDYKEGHF 10

RESULT 10
E53284
T-cell receptor beta 2 chain J region, Jbeta2.2 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: E53284
R:Harinaranath, N.; Alexander, C.B.; Mage, R.G.
Mol. Immunol. 28, 881-888, 1991
A:Title: Evolutionarily conserved organization and sequences of germline diversity and
A:Reference number: A53284; MUID:91342695; PMID:1678859
A:Accession: E53284
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-16 <HAR>
A:Cross-references: GB:S60737; NID:G233916; PIDN:AA19521.1; PID:G233921
A:Note: sequence extracted from NCBI backbone (NCBI:60737, NCBI:60743)
C:Keywords: T-cell receptor

Query Match 25.0%; Score 26; DB 2; Length 16;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 EBYFGD 18
: ||||
Db 4 QLYFGD 9

RESULT 11
H22565
R-phycoerythrin gamma-C chain - red alga (Gastroclonium coulteri) (fragment)
C:Species: Gastroclonium coulteri
C:Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 26-May-2000
C:Accession: H22565
R:Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.

A:Reference number: A22565; MUID:85182601; PMID:3886644
A:Accession: H22565
A:Molecule type: protein
A:Residues: 1-20 <KLO>
C:Superfamily: Aglaothamnion neglectum R-phycoerythrin gamma chain 33

Query Match 25.0%; Score 26; DB 2; Length 20;
Best Local Similarity 42.9%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 LEAKICHQIEEYF 16
: | | | | |
Db 2 LYASSCHHEEQIF 15

RESULT 12
S46550
actin-related protein - chicken (fragments)
C:Species: Gallus gallus (chicken)
C:Date: 12-Sep-1995 #sequence_revision 10-Nov-1995 #text_change 17-Mar-1999
C:Accession: S46550
R:Lees-Miller, J.P.; Helfman, D.M.; Schroer, T.A.
Nature 359, 244-246, 1992
A:Title: A vertebrate actin-related protein is a component of a multisubunit complex
A:Reference number: S29089; MUID:92408781; PMID:1528266
A:Accession: S46550
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-21 <LEE>

Query Match 25.0%; Score 26; DB 2; Length 21;
Best Local Similarity 40.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 HOIEEYFGD 18
: : | | |
Db 12 HRKQYLLXD 21

RESULT 13
I77373
gene N-ras protein - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C:Accession: I77373
R:Kiba, T.; Tsuda, H.; Pairajkul, C.; Inoue, S.; Sugimura, T.; Hirohashi, S.
Mol. Carcinog. 8, 312-318, 1993
A:Title: Mutations of the p53 tumor suppressor gene and the ras gene family in int
A:Reference number: I57484; MUID:94107479; PMID:8280380
A:Accession: I77373
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <RES>
A:Cross-references: GB:S68581; NID:G544858; PIDN:AA29640.1; PID:G544859
C:Genetics:
A:Gene: N-ras

Query Match 25.0%; Score 26; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICHQ 10
: ||||
Db 19 ICHQ 22

RESULT 14
PHI725
Ig heavy chain V region (clone GCC-5) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PHI725
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.

J. Exp. Med. 178, 295-307, 1993
A;Title: Antigen-driven B cell differentiation in vivo.
A;Reference number: PH1675; MUID:93301607; PMID:8315385
A;Accession: PH1725
A;Molecule type: mRNA
A;Residues: 1-23 <MCH>
A;Experimental source: B cell
A;Note: the authors translated the codon ACA for residue 13 as Ala
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotrimer; immunoglobulin

Query Match 25.0%; Score 26; DB 2; Length 23;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 14 YFEGDF 19
||:|:
Db 15 YYVGSY 20

RESULT 15

C57001
endo-1,4-beta-xylanase (EC 3.2.1.8) 4 - Streptomyces roseiscleroticus (fragment)
C;Species: Streptomyces roseiscleroticus
C;Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 23-May-1997
C;Accession: C57001
R;Grabski, A.C.; Forrester, I.T.; Patel, R.; Jeffries, T.W.
Protein Expr. Purif. 4, 120-129, 1993
A;Title: Characterization and N-terminal amino acid sequences of beta-(1-4)endoxylanases
A;Reference number: A57001; MUID:93229899; PMID:8471845
A;Accession: C57001
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-25 <GRA>
A;Experimental source: strain NRRL B-11019
A;Note: sequence extracted from NCBI backbone (NCBIP:130009)
C;Function:
A;Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
A;Pathway: xylan degradation
C;Superfamily: Streptomyces endo-1,4-beta-xylanase A; Streptomyces endo-1,4-beta-xylanase
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation

Query Match 25.0%; Score 26; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 YFEG 17
||||
Db 14 YFEG 17

Search completed: April 23, 2003, 13:48:55
Job time : 10.927 secs

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1

Query Match 85.1%; Score 88.5; DB 9; Length 18;
Best Local Similarity 94.7%; Pred. No. 8.8e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
|||||
Db 1 AALEAKICHQI-EYFGDF 18

RESULT 3
US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: 09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14

Query Match 85.1%; Score 88.5; DB 9; Length 18;
Best Local Similarity 94.7%; Pred. No. 8.8e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
|||||
Db 1 AALEAKICHQI-EYFGDF 18

RESULT 4
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13

Query Match 81.2%; Score 84.5; DB 9; Length 17;
Best Local Similarity 94.4%; Pred. No. 3.6e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2 AALEAKICHQIEYYFGDF 19
|||||
Db 1 AALEAKICHQI-EYFGDF 17

RESULT 5
US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match 80.3%; Score 83.5; DB 9; Length 18;
Best Local Similarity 89.5%; Pred. No. 5.5e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
|||||
Db 1 AALEAKICHQIEQ-YFGDF 18

RESULT 6
US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match 79.3%; Score 82.5; DB 9; Length 18;
Best Local Similarity 94.4%; Pred. No. 8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGD 18
|||||
Db 1 AALEAKICHQI-EYFGD 17

RESULT 7
US-09-836-073-12

```
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match      77.4%; Score 80.5; DB 9; Length 18;
Best Local Similarity 89.5%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
Db 1 AALEAKICHQI-EYQFGDF 18

RESULT 8
US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match      76.4%; Score 79.5; DB 9; Length 18;
Best Local Similarity 89.5%; Pred. No. 2.4e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
Db 1 AALEAKICHQI-EYQFGDF 18

RESULT 9
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 2002-10-24
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2

Query Match      73.6%; Score 76.5; DB 9; Length 18;
Best Local Similarity 84.2%; Pred. No. 7.2e-06;
Matches 16; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
Db 1 AALEAQICQI-EYYFGDF 18

RESULT 10
US-09-836-073-3
; Sequence 3, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-3

Query Match      73.6%; Score 76.5; DB 9; Length 18;
Best Local Similarity 78.9%; Pred. No. 7.2e-06;
Matches 15; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
Db 1 AALQAKICHQI-QYYFGQF 18

RESULT 11
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match      73.6%; Score 76.5; DB 9; Length 18;
Best Local Similarity 93.8%; Pred. No. 7.2e-06;
```

Matches 15; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 4 EAKICHQIEYYFGDF 19
||||| |||||
Db 4 EAKICHQI-EYYFGDF 18

RESULT 12

US-09-836-073-8
; Sequence 8, Application US/09836073
; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Balidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 22002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-8

Query Match 72.6%; Score 75.5; DB 9; Length 18;

Best Local Similarity 84.2%; Pred. No. 1e-05; Indels 1; Gaps 1;

Matches 16; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 AALEAKICHQIEYYFGDF 19

||||| |||||

Db 1 AALEAKICHQIEQ-QFGDF 18

RESULT 13

US-09-836-073-7

; Sequence 7, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Balidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 22002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-7

Query Match 70.7%; Score 73.5; DB 9; Length 18;

Best Local Similarity 88.9%; Pred. No. 2.2e-05; Indels 1; Gaps 1;

Matches 16; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 AALEAKICHQIEYYFGD 18

||||| |||||

Db 1 AALEAKICHQI-EYQGD 17

RESULT 14

US-09-836-073-15

; Sequence 15, Application US/09836073

; Patent No. US20020173475A1

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:38:28 ; Search time 8.96629 seconds
(without alignments)
62.349 Million cell updates/sec

Title: US-09-836-073-16

Perfect score: 104

Sequence: 1 AALEAKICHQIEYYFGDF 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 147762

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88.5	85.1	18	4	US-09-316-630-3
2	88.5	85.1	18	4	US-09-316-630-4
3	31	29.8	20	4	US-08-505-250-8
4	31	29.8	20	4	US-08-505-250-8
5	30	28.8	7	1	US-08-166-930-15
6	30	28.8	7	2	US-08-727-045A-15
7	30	28.8	7	4	US-09-408-172-15
8	30	28.8	11	2	US-08-211-312-9
9	30	28.8	11	3	US-08-472-285-9
10	30	28.8	11	4	US-08-472-929-9
11	29	27.9	13	1	US-08-179-481-91
12	29	27.9	14	4	US-09-298-924-55
13	29	27.9	14	4	US-08-469-260A-455
14	29	27.9	23	2	US-08-808-277A-38
15	29	27.9	23	3	US-08-746-160-45
16	28	26.9	10	2	US-08-618-696-2
17	28	26.9	10	2	US-08-618-696-10
18	28	26.9	10	2	US-08-618-696-11
19	28	26.9	10	3	US-09-033-753-2
20	28	26.9	10	3	US-09-033-753-10
21	28	26.9	10	3	US-09-033-753-11
22	28	26.9	11	2	US-08-618-696-1
23	28	26.9	11	2	US-08-618-696-6
24	28	26.9	11	2	US-08-618-696-7
25	28	26.9	11	2	US-08-618-696-16
26	28	26.9	11	2	US-08-618-696-19
27	28	26.9	11	2	US-08-618-696-20

28	28	26.9	11	3	US-09-033-753-1	Sequence 1, Appli
29	28	26.9	11	3	US-09-033-753-6	Sequence 6, Appli
30	28	26.9	11	3	US-09-033-753-7	Sequence 7, Appli
31	28	26.9	11	3	US-09-033-753-16	Sequence 16, Appli
32	28	26.9	11	3	US-09-033-753-19	Sequence 19, Appli
33	28	26.9	11	3	US-09-033-753-20	Sequence 20, Appli
34	28	26.9	15	4	US-08-743-168B-29	Sequence 29, Appli
35	28	26.9	15	5	PCT-US96-10435-29	Sequence 29, Appli
36	28	26.9	22	6	5281520-43	Patent No. 5281520
37	27	26.0	8	4	US-09-187-859-745	Sequence 745, App
38	27	26.0	8	4	US-09-305-927-155	Sequence 155, App
39	27	26.0	9	4	US-09-077-977A-7	Sequence 7, Appli
40	27	26.0	9	4	US-09-187-859-746	Sequence 746, App
41	27	26.0	9	4	US-09-305-927-156	Sequence 156, App
42	27	26.0	10	4	US-09-187-859-747	Sequence 747, App
43	27	26.0	10	4	US-09-305-927-157	Sequence 157, App
44	27	26.0	11	2	US-08-641-314C-6	Sequence 6, Appli
45	27	26.0	12	1	US-07-894-212A-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match 85.1%; Score 88.5; DB 4; Length 18;
Best Local Similarity 94.7%; Pred. No. 3.7e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEYYFGDF 19
|||||
Db 1 AALEAKICHQI-EYYFGDF 18

RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
; OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 85.1%; Score 88.5; DB 4; Length 18;
Best Local Similarity 94.7%; Pred. No. 3.7e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEEYFGDF 19

Db 1 AALEAKICHQI-EYFGDF 18

RESULT 3
US-08-505-250-8
; Sequence 8, Application US/08505250
; Patent No. 6183983
; GENERAL INFORMATION:
; APPLICANT: Sato, Haruya
; APPLICANT: Yamamoto, Keiji
; APPLICANT: Suzuki, Kokichi
; APPLICANT: Ikeda, Masahiro
; APPLICANT: Sakagami, Masahiro
; APPLICANT: Taniguchi, Makoto
; TITLE OF INVENTION: PROTEIN MODIFICATION METHOD
; FILE REFERENCE: 110-511
; CURRENT APPLICATION NUMBER: US/08/505,250
; CURRENT FILING DATE: 1995-11-29
; EARLIER APPLICATION NUMBER: PCT/JP95/00298
; EARLIER FILING DATE: 1995-02-27
; EARLIER APPLICATION NUMBER: JP 198187/94
; EARLIER FILING DATE: 1994-08-23
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-08-505-250-8

Query Match 29.8%; Score 31; DB 4; Length 20;
Best Local Similarity 44.4%; Pred. No. 87;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 KICHOIEEY 14

Db 3 KMCPQLQY 11

RESULT 4
US-08-505-250-8
; Sequence 8, Application US/08505250
; Patent No. 6322996
; GENERAL INFORMATION:
; APPLICANT: Sato, Haruya
; APPLICANT: Yamamoto, Keiji
; APPLICANT: Suzuki, Kokichi
; APPLICANT: Ikeda, Masahiro

; APPLICANT: Sakagami, Masahiro
; APPLICANT: Taniguchi, Makoto
; TITLE OF INVENTION: PROTEIN MODIFICATION METHOD
; FILE REFERENCE: 110-511
; CURRENT APPLICATION NUMBER: US/08/505,250
; CURRENT FILING DATE: 1995-11-29
; PRIOR APPLICATION NUMBER: PCT/JP95/00298
; PRIOR FILING DATE: 1995-02-27
; PRIOR APPLICATION NUMBER: JP 198187/94
; PRIOR FILING DATE: 1994-08-23
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
US-08-505-250-8

Query Match 29.8%; Score 31; DB 4; Length 20;
Best Local Similarity 44.4%; Pred. No. 87;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 6 KICHOIEEY 14

Db 3 KMCPQLQY 11

RESULT 5
US-08-166-930-15
; Sequence 15, Application US/08166930
; Patent No. 5599678
; GENERAL INFORMATION:
; APPLICANT: Kraus, Michael
; APPLICANT: Stuber, Werner
; TITLE OF INVENTION: Synthetic Peptides, Antibodies Against
; TITLE OF INVENTION: Them and Their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/166,930
; FILING DATE: 15-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 42 42 736.3
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Einandi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 02481.1351-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-08-166-930-15

Query Match 28.8%; Score 30; DB 1; Length 7;
Best Local Similarity 80.0%; Pred. No. 2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQIE 12
|||:|
Db 1 CHQVE 5

RESULT 6

US-08-727-045A-15
; Sequence 15, Application US/08727045A
; Patent No. 5981697
; GENERAL INFORMATION:
; APPLICANT: Kraus, Michael
; APPLICANT: Stuber, Werner
; TITLE OF INVENTION: Synthetic Peptides, Antibodies Against
; TITLE OF INVENTION: Them and Their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/727,045A
; FILING DATE: 08-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 42 42 736.3
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 05552.1351-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-727-045A-15

Query Match 28.8%; Score 30; DB 2; Length 7;
Best Local Similarity 80.0%; Pred. No. 2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQIE 12
|||:|
Db 1 CHQVE 5

RESULT 7

US-09-408-172-15
; Sequence 15, Application US/09408172
; Patent No. 6441141
; GENERAL INFORMATION:
; APPLICANT: Kraus, Michael
; APPLICANT: Stuber, Werner
; TITLE OF INVENTION: Synthetic Peptides, Antibodies Against

; TITLE OF INVENTION: Them and Their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/408,172
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/727,045
; FILING DATE: 08-OCT-1996
; APPLICATION NUMBER: DE P 42 42 736.3
; FILING DATE: 17-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 05552.1351-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-408-172-15

Query Match 28.8%; Score 30; DB 4; Length 7;
Best Local Similarity 80.0%; Pred. No. 2e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 CHQIE 12
|||:|
Db 1 CHQVE 5

RESULT 8

US-08-211-312-9
; Sequence 9, Application US/08211312
; Patent No. 5986051
; GENERAL INFORMATION:
; APPLICANT: LABIGNE, AGNES
; APPLICANT: CUSSAC, VALERIE
; APPLICANT: FERRERO, RICHARD
; TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
; TITLE OF INVENTION: FOR THE REGULATION AND MATURATION OF UREASE AND THEIR U
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

; REFERENCE/DOCKET NUMBER: 660-075-0XPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEX: (703) 413-2220
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-472-285-9

Query Match 28.8%; Score 30; DB 3; Length 11;
Best local Similarity 71.4%; Pred. No. 67;
Matches 5; Conservative 2; Mismatches 0; Indels

QY 5 AKICHQI 11
DB 1 AKICYEI 7

RESULT 10
US-08-472-929-9
; Sequence 9, Application US/08472929
; Patent No. 6271017
; GENERAL INFORMATION:
; APPLICANT: LABIGNE, AGNES
; APPLICANT: CUSSAC, VALERIE
; APPLICANT: FERRERO, RICHARD
; TITLE OF INVENTION: GENES OF HELICOBACTER PYLORI NECESSARY
; TITLE OF INVENTION: FOR THE REGULATION AND MATURATION OF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBOLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,929
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/211,312
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR92/00921
; FILING DATE: 02-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Obolon, No. 6271017man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-075-0XPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEX: (703) 413-2220
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-472-929-9

Query Match 28.8%; Score 30; DB 4; Length 11;

```

Best Local Similarity 71.4%; Pred. No. 67;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 AKICHOI 11
Db 1 AKICVEI 7

RESULT 11

US-08-179-481-91
; Sequence 91, Application US/08179481
; Patent No. 5624816
; GENERAL INFORMATION:
; APPLICANT: CARRAWAY, KERMIT L.
; APPLICANT: CARRAWAY, CORALIE A.
; APPLICANT: FREGEN, NEVIS L.
; TITLE OF INVENTION: ONCOGENE PRODUCT LIGAND
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/179,481
; FILING DATE: 28-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/922,521
; FILING DATE: 30-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 200702/UM92-08CIP
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-179-481-91

Query Match 27.9%; Score 29; DB 1; Length 13;
Best Local Similarity 46.2%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 LEAKICHIEEY 15
Db 1 LEAVTCKQOEYF 13

RESULT 12

US-09-298-924-55
; Sequence 55, Application US/09298924
; Patent No. 6391595
; GENERAL INFORMATION:
; APPLICANT: KATO, Masaru
; MIURA, Yutaka
; KETTOKU, Masako
; IWAMATSU, Akihiro
; KOBAYASHI, Kazuo

KOMEDA, Toshihiro
; TITLE OF INVENTION: NOVEL TRANSFERASE AND AMYLASE, PROCESS
; FOR PRODUCING THE ENZYMES, USE THEREOF, AND GENE CODING
; FOR THE SAME
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/298,924
; FILING DATE: 26-Apr-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/750,569
; FILING DATE: <Unknown>
; APPLICATION NUMBER: JP 7-120673
; FILING DATE: 21-APR-1995
; APPLICATION NUMBER: JP 6-311185
; FILING DATE: 21-NOV-1994
; APPLICATION NUMBER: JP 6-286917
; FILING DATE: 21-NOV-1994
; APPLICATION NUMBER: JP 6-290394
; FILING DATE: 31-OCT-1994
; APPLICATION NUMBER: JP 6-194223
; FILING DATE: 18-AUG-1994
; APPLICATION NUMBER: JP 6-133354
; FILING DATE: 16-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 49441/110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-298-924-55

Query Match 27.9%; Score 29; DB 4; Length 14;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 13 EYFGDF 19
Db 1 DYYYQDF 7

RESULT 13

US-08-469-260A-455
; Sequence 455, Application US/08469260A
; Patent No. 6451578
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY

;; APPLICANT: ANTHONY SCOTT MUERHOFF
;; APPLICANT: JAMES C. ERKER
;; APPLICANT: SHERI L. BUIJK
;; APPLICANT: ISA K. MUSHAWAR
;; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
;; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
;; NUMBER OF SEQUENCES: 716
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
;; STREET: 100 ABBOTT PARK ROAD
;; CITY: ABBOTT PARK
;; STATE: IL
;; COUNTRY: USA
;; ZIP: 60064-3500
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/469,260A
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/424,550
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FOREMSKI, PRISCILLA E.
;; REGISTRATION NUMBER: 33,207
;; REFERENCE/DOCKET NUMBER: 5527.PC.01
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 708-937-6365
;; TELEFAX: 708-938-2623
;; INFORMATION FOR SEQ ID NO: 455;
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 14 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-469-260A-455

Query Match 27.9%; Score 29; DB 4; Length 14;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 AKICQH 10
| | | | |
Db 3 AMICQH 8

RESULT 14
US-08-808-277A-38
; Sequence 38, Application US/0808277A
; Patent No. 5993374
; GENERAL INFORMATION:
; APPLICANT: LEHRER, ROBERT I.
; APPLICANT: LEE, IN-HEE
; APPLICANT: ZHAO, CHENGQUAN
; TITLE OF INVENTION: CLAVASPIRINS
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/746,160
; FILING DATE: 06-NOV-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20563.00
; TELECOMMUNICATION INFORMATION:
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,277A
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 220002057300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-822-0168
; TELEX:

;; APPLICATION NUMBER: US/08/808,277A
;; FILING DATE: 28-FEB-1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Murashige, Kate H
;; REGISTRATION NUMBER: 29,959
;; REFERENCE/DOCKET NUMBER: 220002057300
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-887-1500
;; TELEFAX: 202-822-0168
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 38;
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 23 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; FEATURE:
;; NAME/KEY: Modified Base
;; LOCATION: 17..0
;; OTHER INFORMATION: o-methyl tyrosine
;; US-08-808-277A-38

Query Match 27.9%; Score 29; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 LEAKICHOIEEYFG 17
| | | | |
Db 4 LLGKIIHHVGNFVYG 18

RESULT 15
US-08-746-160-45
; Sequence 45, Application US/08746160
; Patent No. 6010876
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Harwig, Sylvia L.
; APPLICANT: Zhao, Chengquan
; APPLICANT: Lee, In-Hee
; TITLE OF INVENTION: CLAVANINS
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/746,160
; FILING DATE: 06-NOV-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20563.00
; TELECOMMUNICATION INFORMATION:
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,277A
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 220002057300
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-822-0168
; TELEX:

```

; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 23 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: No. 6010876e
;   FEATURE:
;     NAME/KEY: Other
;     LOCATION: 17...17
;     OTHER INFORMATION: o-methyl tyrosine
US-08-746-160-45

Query Match      27.9%; Score 29; DB 3; Length 23;
Best Local Similarity 33.3%; Pred. NO. 2.le+02;
Matches 5; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      3 LEAKICHQIEYYFG 17
      | | | | : : |
Db      4 LLGKIHHVGNFYVG 18

```

Search completed: April 23, 2003, 13:50:27
Job time : 9.96629 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:28:18 ; Search time 28.7135 Seconds
(without alignments)
88.173 Million cell updates/sec

Title: US-09-836-073-16

Perfect score: 104
Sequence: 1 AALEAKICHQIEYYRGDF 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 320064

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	88.5	85.1	18	AAV52200	Human la autoantigen
2	62.5	60.1	21	AA843394	La/SSB epitope 17.
3	43	41.3	23	AB879195	Human prostate spe
4	35	33.7	21	AA892191	Signal transductio
5	35	33.7	24	AAV94731	Region B of protei
6	35	33.7	24	AA899801	Protein kinase der
7	34	32.7	22	AAW00599	Heparan sulphate p
8	33	31.7	16	AAW98909	Vaccine related MH
9	33	31.7	25	AAW90166	Human c-FLIP-L pep
10	33	31.7	25	AA803965	FLICE like inhibit

11	32	30.8	9	20	AAV10410	T cell epitope/MHC
12	32	30.8	18	23	AAE20865	E7-novispirin pept
13	31	29.8	8	14	AA843395	La/SSB epitope 24.
14	31	29.8	10	23	AB884046	Transglutaminase i
15	31	29.8	10	23	AB822202	Murine MC-1 antibo
16	30.5	29.3	22	20	AAV19621	SEQ ID NO 339 from
17	30	28.8	7	15	AA856273	Peptide antigen us
18	30	28.8	10	21	AAV66804	r cell antigen rec
19	30	28.8	11	14	AA834400	Fragment of Helico
20	30	28.8	14	22	AA897583	Human peptide #858
21	30	28.8	15	22	AB852123	Human API-202 tryp
22	30	28.8	18	22	AB89021	HIV gp120 protein
23	30	28.8	18	22	AB89022	HIV gp120 protein
24	30	28.8	20	21	AAV87761	Potato THT protein
25	30	28.8	20	21	AAV87765	Potato THT protein
26	30	28.8	20	23	AAU90254	Insulin/insulin-11
27	30	28.8	24	22	AA893687	Anti-Rh(D) heavy c
28	29	27.9	8	16	AA873216	Human TSH receptor
29	29	27.9	9	19	AAW72483	Dengue virus type-
30	29	27.9	14	17	AA890665	S. acidocaldarius
31	29	27.9	14	21	AA809328	Hepatitis GB virus
32	29	27.9	18	22	AA820426	Anti-FIX/FIXa anti
33	29	27.9	18	23	AA820866	D7-novispirin pept
34	29	27.9	20	15	AA861169	Peptide for treati
35	29	27.9	20	18	AAW34602	Partial sequence o
36	29	27.9	21	19	AAW63669	Antimicrobial pept
37	29	27.9	21	22	AA873435	Native human TSHR
38	29	27.9	21	23	AAU89799	Insulin/insulin-11
39	29	27.9	23	19	AAW63664	Antimicrobial pept
40	28	26.9	9	23	AB847453	N. meningitidis LO
41	28	26.9	10	15	AA856293	Native human neutr
42	28	26.9	10	15	AA856296	Synthetic modified
43	28	26.9	10	15	AA856297	Synthetic modified
44	28	26.9	11	15	AA856306	Modified Influenza
45	28	26.9	11	15	AA856307	Modified Influenza

ALIGNMENTS

RESULT 1
AAV52200
ID AAV52200 standard; peptide; 18 AA.
AC AAV52200;

DT 14-MAR-2000 (first entry)
XX Human la autoantigen peptide (LAP).

DE La autoantigen; LAP; Internal ribosome entry site; IRES; translation;
KW viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
KW coronavirus; hepatitis virus; rhadovirus; adenovirus; coxsackie virus;
KW parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
KW vesicular stomatitis virus.

OS Homo sapiens.

XX WO9961613-A2.

XX 02-DEC-1999.

XX 21-MAY-1999; 99WO-US11281.

XX 22-MAY-1998; 98US-0086527.

PA (REGC) UNIV CALIFORNIA.

PI Das S, Dasgupta A;

XX WPI; 2000-062712/05.

XX

PT New yeast inhibitory peptide useful for inhibiting viral protein
 XX translation and replication -
 PS Claim 5; Page 57; 81pp; English.
 XX This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (IRNA see AA245200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rhinovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.
 XX
 SQ Sequence 18 AA;

Query Match 85.1%; Score 88.5; DB 21; Length 18;
 Best Local Similarity 94.7%; Pred. No. 2.4e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AALEAKICHQIEEYVFGDF 19
 DB 1 AALEAKICHQI-EYVFGDF 18
 |||||

RESULT 2
 AAR43394
 ID AAR43394 standard; peptide; 21 AA.
 XX
 AC AAR43394;

DT 12-MAY-1994 (first entry)

DE La/SSB epitope 17.

KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

OS Homo sapiens.

PN WO9321223-A.

XX 28-OCT-1993.

PF 13-APR-1993; 93WO-US03484.

XX 13-APR-1992; 92US-0867819.

PR (OKLA) UNIV OKLAHOMA STATE.

XX Harley JB;

DR WPI; 1993-351658/44.

XX New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus

PS Claim 1; Page 30; 43pp; English.

XX The sequences given in AAR43391-562 are linear epitopes which are

CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B',
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polypyridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides; B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.
 XX
 SQ Sequence 21 AA;

Query Match 60.1%; Score 62.5; DB 14; Length 21;
 Best Local Similarity 92.3%; Pred. No. 0.004;
 Matches 12; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 7 ICHQIEEYVFGDF 19
 DB 1 ICHQI-EYVFGDF 12
 |||||

RESULT 3
 ABB79195
 ID ABB79195 standard; Protein; 23 AA.
 XX
 AC ABB79195;

DT 08-AUG-2002 (first entry)

XX Human prostate specific protein sequence SEQ ID NO:144.

DE Human; prostate specific gene; prostate specific protein; PSG; PSP;
 KW prostate cancer.

XX Homo sapiens.

XX WO200236808-A2.

PN 10-MAY-2002.

XX 05-NOV-2001; 2001WO-US47283.

XX 03-NOV-2000; 2000US-245740P.

XX (DIAD-) DIADEXUS INC.

XX Sun Y, Recipon H, Chen S, Liu C;

XX WPI; 2002-471506/50.

XX New prostate-specific nucleic acids and polypeptides, useful for
 PT identifying, diagnosing, monitoring, staging, imaging, and treating
 PT prostate cancer and non-cancerous disease states in prostate tissue

PS Claim 11; Page 218; 254pp; English.

XX ABN87650 to ABN87789 represent human prostate-specific nucleic acids (I),
 CC and ABB79192 to ABB79295 represent human prostate-specific proteins (II)
 CC from the present invention. (I) and (II) have cytostatic activity. (I)
 CC can be used in gene therapy. The prostate-specific nucleic acids,
 CC polypeptides and compositions from the present invention can be used for
 CC identifying, diagnosing, monitoring, staging, imaging, and treating
 CC prostate cancer and non-cancerous disease states in prostate tissue; for

CC identifying prostate tissue; for monitoring, identifying and/or designing
 CC agonists and antagonists of the polypeptides; in gene therapy; in
 CC producing transgenic animals and cells; for producing engineered prostate
 CC tissue for treatment and research; and as elements in an array or
 CC computer program for pattern recognition of prostate disorders. The
 CC nucleic acids may be used as hybridisation probes to detect, characterise
 CC and quantify hybridising nucleic acids in, and isolate hybridising
 CC nucleic acids from, both genomic and transcript-derived nucleic acid
 CC samples.
 CC
 SQ Sequence 23 AA;

Query Match 41.3%; Score 43; DB 23; Length 23;
 Best Local Similarity 46.7%; Pred. No. 5.7;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 ALEAKICHQIEYYF 16
 II: III :II
 Db 6 ALQTPICHTGKEVF 20

RESULT 4
 AAB92191
 ID AAB92191 standard; Peptide; 21 AA.

XX AAB92191;
 XX
 DT 22-JUN-2001 (first entry)
 XX
 DE Signal transduction reagents peptide SEQ ID NO:1367.
 XX
 KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
 KW blood component; modification; succinimidyl; maleimido group; amino;
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
 XX
 OS Homo sapiens.
 OS Synthetic.

XX WO200069900-A2.
 XX 23-NOV-2000.
 XX 17-MAY-2000; 2000WO-US13576.
 XX 17-MAY-1999; 99US-0134406.
 PR 10-SEP-1999; 99US-0153406.
 PR 15-OCT-1999; 99US-0159783.
 XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
 PI WPI; 2001-112059/12.
 DR
 XX
 PT Modifying and attaching therapeutic peptides to albumin prevents
 PT peptidase degradation, useful for increasing length of in vivo activity
 PT
 XX
 PS Disclosure; Page 643; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)
 CC comprising a therapeutically active amino acid region (III) and a
 CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
 CC a less therapeutically active amino acid region (IV), which covalently
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a
 CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
 CC factors and neurotransmitters, to protect them from peptidase activity
 CC in vivo for the treatment of various disorders. Endogenous therapeutic
 CC peptides are not suitable as drug candidates as they require frequent
 CC administration due to rapid degradation by peptidases in the body.
 CC Modifying and attaching therapeutic peptides to albumin prevents or
 CC reduces the action of peptidases to increase length of activity (half

CC life) and specificity as bonding to large molecules decreases
 CC intracellular uptake and interference with physiological processes.
 CC AAB90829 to AAB92441 represent peptides which can be used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 21 AA;

Query Match 33.7%; Score 35; DB 22; Length 21;
 Best Local Similarity 75.0%; Pred. NO. 99;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 QIEEYVFG 17
 I: IIII
 Db 8 QAEWVFG 15

RESULT 5
 AAY94731
 ID AAY94731 standard; peptide; 24 AA.

XX AAY94731;
 XX
 DT 29-JAN-2001 (first entry)
 XX
 DE Region B of protein kinase C related protein kinase 2.
 XX
 KW Substrate specificity; phosphoinositide-dependent protein kinase 1;
 KW PDK1; protein kinase C related protein kinase 2; PRK2; cancer; apoptosis;
 KW mechanical tissue damage; ischaemic disease; stroke;
 KW myocardial infarction; antigenic peptide.
 XX
 OS Unidentified.

XX WO200056864-A2.
 XX 28-SEP-2000.
 XX 17-MAR-2000; 2000WO-GB01004.
 PR 19-MAR-1999; 99GB-0006245.
 XX (UYDU-) UNIV DUNDEE.
 XX
 PI Alessi D, Balendran A, Deak M, Currie R, Downes P, Casamayor A;
 XX WPI; 2000-647155/62.

XX Altering substrate specificity of phosphoinositide-dependent protein
 XX kinase 1, to phosphorylate Ser473 in addition to Thr308 by exposing to
 XX interacting polypeptide -
 XX
 PS Disclosure; Page 12; 103pp; English.

XX This invention relates to a method for altering the substrate
 CC specificity of phosphoinositide-dependent protein kinase 1 (PDK1), by
 CC exposing it to an interacting polypeptide. Included in the invention are
 CC a preparation comprising PDK1 and an interacting polypeptide, PDK1 with
 CC altered specificity is useful for phosphorylating a residue corresponding
 CC to the Ser/Thr residue of a substrate with the following peptide
 CC Phe/Tyr-xaa-xaa-Phe/Tyr-Ser/Thr-Phe/Tyr. Altered PDK1 is also used for
 CC phosphorylating protein kinase C related protein kinase 2 (PRK2). The
 CC compound identified by methods of the invention that are capable of
 CC altering the substrate specificity of PDK1 are useful for manufacturing a
 CC medicament for treating a patient who is in need of modulation of the
 CC insulin signalling pathway and/or PDK1, PRK2 or PRK2 signalling. A
 CC compound that is capable of reducing the activity (i.e. the PDK1 and/or
 CC the PDK2 activity) of PDK1 may be useful in treating cancer. PDK1,
 CC e.g. via protein kinase B and/or SGK, may be capable of providing a
 CC survival signal that protects cells from apoptosis induced in a variety
 CC of ways. Reduction of the activity of PDK1 may promote apoptosis and may
 CC be useful in treating cancer. Conditions in which aiding apoptosis may be
 CC of benefit may also include resolution of inflammation. A compound
 CC capable of increasing the activity of PDK1 may be useful in treating

composition is also useful for inhibiting the degree or rate of phosphorylation by the protein kinase. The interacting polypeptide or compound is useful in methods of stabilising a hydrophobic pocket-containing protein kinase, where the protein kinase is exposed to the compound or polypeptide. AAB93786 to AAB99847 represent amino acid sequences, and AAH44210 and AAH44211 represent oligonucleotide sequences, used in the exemplification of the present invention.

```
Query Match      33.7%; Score 35; DB 22; Length 24;
Best Local Similarity 43.8%; Pred. NO. 1.le+02;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

QY 4 EAKICHQIEEYYFGDF 19
 | : | : | : |
Db 2 EPRILSEEEQEEMFGDF 17

RESULT 7
AAW00599
ID AAW00599 standard; peptide; 22 AA.

DT 24-MAR-1997 (first entry)

DE Heparan sulphate proteoglycan core protein fragment v18.

KW Heparan sulphate proteoglycan; core protein; HSPG; V8 digestion; murine;
 KW tryptic peptide; neuroepithelial cell; c-myc; oncogene; embryonic cell;
 KW glycosaminoglycan; GAG; fibroblast growth factor; FGF; promoter; mouse;
 KW cytokine; cell proliferation; motor neuronal cell; neurological disease;
 KW FGF type IIIc receptor; Huntington's disease; Parkinson's disease;
 KW therapy.

OS Mus musculus.

PN WO9623003-A1.

25-JAN-1996; 96WO-AU000034.

16-JUN-1995: 95AU-0003560.

27 JAN 1955, 0540 0000/04:
XX
PA (AMRA-) AMRAD OPERATIONS PTY LTD.

XX
XX
FBI / BUREAU OF INVESTIGATION

PI Bartlett PF, Nurcombe V;

XX Glycosaminoglycan polymer or fragment that interacts with cytokine,
PT esp. FGF - used to rescue neuronal cells e.g. in transplants of
PT FGF-expressing cells for treatment of neurological disease
XX WFL, 1990-302030/36.

Example 1: page 21: 57pp: English:

AAW00599 and AAW00600 represent fragments of the heparan sulphate proteoglycan (HSPG) core protein. This sequence was generated from v8 digestion of HSPG which was produced by the murine neuroepithelial cell line 2.3D. The 2.3D cell line is made by expressing the c-myc oncogene in cloned embryonic day 10 primary neuroepithelial cells. HSPG is used to prepare the glycosaminoglycan (GAG) polymers of the invention. The GAGs of the invention are able to interact with fibroblast growth factor-1 (FGF-1) or FGF-2, but not both. They are used to promote interaction between a specific cytokine and its target site on an animal cell. The GAGs can also be used to promote cell proliferation, migration, and/or differentiation of any tissue bearing the appropriate FGF receptor. The GAGs promote maintenance/survival of motor neuronal cells, and the viability of cells with the FGF type IIIc receptor. A particular application of the GAGs is to treat a neurological disease.

CC (especially Huntington's or Parkinson's), in conjunction with
 CC transplantation of cells that express FGF. Fragments of GAG that are
 CC able to bind to FGF have antagonistic activity, so that they prevent the
 CC formation of an activating FGF-heparan sulphate-FGF receptor ternary
 CC complex. These HSPG core sequences are useful as substrates for
 CC synthesis of GAG, or for the production of antibodies, antagonists and
 CC agonists of GAG. Hybrids of GAG with a specific targeting peptide can
 CC be used for localised treatment.

XX
 SQ Sequence 22 AA;

Query Match 32.7%; Score 34; DB 17; Length 22;
 Best Local Similarity 36.4%; Pred. No. 1.5e+02;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 CHQIEEYFGD 18
 | : : : :
 Db 4 CEDQTFYGD 14

RESULT 8
 AAM98909
 ID AAM98909 standard; Peptide; 16 AA.

XX AC AAM98909;
 XX DT 07-DEC-2001 (first entry)
 XX DE Vaccine related MHC ligand peptide SEQ ID NO:12.

XX KW Glutamic acid; glutamine; vaccine; major histocompatibility complex;
 KW MHC; immunomodulator; antiallergic; endocrine; neuroprotectant;
 KW virucidal; bactericidal; antiparasitic; fungicidal; cytostatic;
 KW medicine; pharmaceutical; immune disorder; immune deficiency;
 KW autoimmune; hypersensitivity; allergy; graft rejection; infection;
 KW hormonal disorder; central nervous system disease; cancer; melanoma;
 KW anti-melanoma vaccine; human immunodeficiency virus.

XX OS Rabies virus.
 XX PN WO200170772-A2.

XX PD 27-SEP-2001.

XX PF 22-MAR-2001; 2001WO-FR00872.

XX PR 23-MAR-2000; 2000FR-0003711.

XX PA (FABR) FABRE MEDICAMENT SA PIERRE.

XX PI Klinguer-Hamour C, Corvaia N, Beck A, Goetsch L;

XX PS WPI; 2001-611470/70.

XX PT Stabilized pharmaceutical containing N-terminal glutamic acid or
 PT glutamine, useful e.g. in anti-melanoma vaccines, is an addition salt
 PT with strong acid -

XX PS Claim 9; Page 31; 149pp; French.

XX CC The present invention describes a pharmaceutical compound (I) that
 CC contains an N-terminal glutamic acid (Glu) or glutamine (Gln) residue
 CC in the form of an addition salt with a strong, physiologically
 CC acceptable acid (II). Also described are: (a) a pharmaceutical
 CC composition containing at least one (I); (b) a vaccine containing at
 CC least one (I) where this is a major histocompatibility complex (MHC)
 CC ligand (Ia); (c) a method for in vitro diagnosis of diseases associated
 CC with the presence of (Ia); (d) a kit for method (c) that includes a (Ia);
 CC and (e) a process for preparing (I). (I) has immunomodulator, endocrine,
 CC antiallergic, neuroprotectant, virucidal, bactericidal, antiparasitic,
 CC fungicidal and cytostatic activities. (I) are useful, in human or
 CC veterinary medicine, in pharmaceutical compositions (for treating immune
 CC disorders, e.g. immune deficiency, autoimmune states, hypersensitivity,

CC allergy, graft rejection, infection, hormonal disorders and central
 CC nervous system diseases), also, where (I) is a MHC ligand (Ia), in
 CC vaccines for treatment or prevention of: (i) viral, bacterial, parasitic
 CC or fungal infections; or (ii) of cancers. A particular application is in
 CC anti-melanoma vaccines. (I) are also useful for in vitro diagnosis of
 CC diseases associated with interactions between MHC and (I). e.g. melanoma
 CC and human immunodeficiency virus infection. AAM98898 to AAM99592
 CC represent peptides which can be used in pharmaceutical compounds from
 CC the present invention.

XX SQ Sequence 16 AA;

Query Match 31.7%; Score 33; DB 22; Length 16;
 Best Local Similarity 60.0%; Pred. No. 1.5e+02;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 ALPAKICHQI 11
 : : : : :
 Db 6 SVEAEIAHQI 15

RESULT 9
 AAW90166
 ID AAW90166 standard; peptide; 25 AA.

XX AC AAW90166;

XX DT 15-MAR-1999 (first entry)

XX DE Human c-FLIP-L peptide epitope.

XX KW Epitope; FLIP protein; cFLIP-L; human; antibody; monoclonal; inhibition;
 KW apoptotic signal transduction; detection; antitumour agent; activator;
 KW immortal; apoptosis-sensitive cell;

XX OS Homo sapiens.

XX PN DE19725847-A1.

XX PD 24-DEC-1998.

XX PF 18-JUN-1997; 97DE-1025847.

XX PR 18-JUN-1997; 97DE-1025847.

XX PA (TSCH/) TSCHOPP J.

XX PA (APOT-) APOTECH SA.

XX PI Bodmer J, Burns K, French E, Hahne M, Hoffmann K;

XX PI Imler M, Rimoldi D, Schneider P, Schroeter M, Steiner V;

XX PI Thome M, Tschoep J, French EL;

XX DR WPI; 1999-061508/06.

XX PT New antibody to FLIP protein - used to suppress inhibition of
 PT apoptotic signal transduction by FLIP proteins, to detect FLIP
 PT proteins and to screen for substances that activate FLIP expression

XX PS Claim 8; Column 17; 14pp; German.

XX CC This sequence represents a human epitope from a FLIP protein, cFLIP-L
 CC which is used to generate antibodies. Such antibodies are used to
 CC suppress the inhibition of apoptotic signal transduction by FLIP
 CC proteins. A monoclonal antibody can be used to detect the corresponding
 CC FLIP protein, to determine the activity of substances that inhibit FLIP
 CC expression (used as antitumour agents), to screen for substances that
 CC activate FLIP expression (e.g. substances that can be used to
 CC immortalise apoptosis-sensitive cells) and to purify the corresponding
 CC FLIP protein.

XX SQ Sequence 25 AA;

Query Match 31.7%; Score 33; DB 20; Length 25;

Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 AKICHQIEE 13
|::|::|
Db 2 AEVIHQVEE 10

RESULT 10

AA03965
ID AAB03965 standard; Protein; 25 AA.

XX AC AAB03965;

DT 26-FEB-2001 (first entry)

DE FLICE like inhibitor protein peptide fragment.

XX Chimeric protein; fusion protein; FLICE like inhibitor protein;
KW FLIP; Fas; TNF; apoptosis; caspase-8; ligand; T cell; thymocyte;
KW tumour specific antigen; immune response; therapy; prophylaxis;
KW diagnosis; HIV; human immunodeficiency syndrome; AIDS;
KW acquired immune deficiency syndrome; human.

XX Homo sapiens.

XX WO200059935-A1.

XX 12-OCT-2000.

XX 05-APR-2000; 2000WO-US09002.

XX 05-APR-1999; 99US-0127867.

XX 06-APR-1999; 99US-0128021.

XX (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.

PA (PAYA/) PAYA C.

PA (ALGE/) ALGECIRAS-SCHMINICH A.

PI Paya C, Algeciras-schminich A;

XX WPI; 2000-664988/64.

XX Fusion polypeptide useful for inhibiting ligand-induced apoptosis,
PT comprises portion of anti-apoptotic polypeptide linked to a transport
PT group

PS Example 1; Page 43; 89pp; English.

XX A chimeric group or fusion peptide which comprises a portion of an
CC anti-apoptotic polypeptide which inhibits apoptosis of lymphocytes in
CC combination with a transport group is described. The transport group
CC is capable of transporting the chimeric group or fusion peptide
CC across the cell membrane. The anti-apoptotic polypeptide is FLICE-like
CC inhibitor protein (FLIP) which inhibits Fas and TNF mediated apoptosis
CC by inhibiting binding of Caspase-8 to the Fas receptor complex, thus
CC shutting off the downstream Fas signalling pathway. The chimeric group
CC and fusion peptide are useful for inhibiting ligand-induced apoptosis
CC by bringing them into contact with T cells. The chimeric group is
CC useful for expanding T cells in vitro e.g. T cells specific for
CC particular antigens such as tumour-specific antigen, for enhancing
CC immune response and to inhibit the apoptosis of chronically activated
CC T cells e.g. activated CD4⁺ T cells in HIV infected patients. The
CC chimeric group is also useful for therapeutic, prophylactic or
CC diagnosis of intracellular delivery of small molecules and
CC macromolecules such as anti-apoptotic polypeptides and nucleic
CC acids encoding such polypeptides. This peptide fragment of FLIP
CC corresponding to amino acids 2-26 was used to inject rats to create
CC anti-human FLIP antiserum.

XX Sequence 25 AA;

Query Match 31.7%; Score 33; DB 21; Length 25;

Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 AKICHQIEE 13
|::|::|
Db 2 AEVIHQVEE 10

RESULT 11

AA10410
ID AAY10410 standard; Peptide; 9 AA.

XX AC AAY10410;

DT 12-MAY-1999 (first entry)

DE T cell epitope/MHC ligand SEQ ID NO:340.

XX Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;
KW immunisation; tumour; infectious disease; immunotherapy; cancer;
KW malignant melanoma; viral disease; hepatitis; AIDS.

XX Synthetic.

OS Simian virus.

PN WO9902183-A2.

XX 21-JAN-1999.

XX 10-JUL-1998; 98WO-US14289.

XX 10-DEC-1997; 97US-0988320.

XX 10-JUL-1997; 97CA-2209815.

XX (CTLI-) CTL IMMUNOTHERAPIES CORP.

XX Kuendig TM, Simard JUL;

XX WPI; 1999-120514/10.

XX Inducing a cytotoxic T lymphocyte response - by maintaining a level
PT of antigen in the lymphatic system of a mammal so as to provide a
PT sustained CTL response, used to treat, e.g. AIDS

XX Disclosure; Page 37; 199pp; English.

XX The present invention describes a method of inducing and/or sustaining
CC an immunological cytotoxic T lymphocyte (CTL) response in a mammal. The
CC method comprises: (a) delivering an antigen to the mammal at a level to
CC induce an immunological CTL response in the mammal; and (b) maintaining
CC the level of the antigen in the mammal's lymphatic system to maintain
CC the immunologic CTL response. The method can be used for the delivery of
CC e.g. a differentiation antigen, a tumour-specific multilinear antigen,
CC an embryonic antigen, an oncogene antigen, a mutated tumour-suppressor
CC gene antigen, or a viral antigen. They can be used for the treatment of
CC disease such as cancer, e.g. malignant melanoma or infectious disease,
CC e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery
CC to the lymphatic system provides for potent CTL stimulation that takes
CC place in the milieu of the lymphoid organ, and it sustains stimulation
CC that is necessary to keep CTL active, cytotoxic and recirculating
CC through the body. AAY10071 to AAY10639 represent examples of peptide
CC antigens given in the present invention.

XX Sequence 9 AA;

Query Match 30.8%; Score 32; DB 20; Length 9;

Best Local Similarity 66.7%; Pred. No. 7.8e+05;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 LEAKICHOI 11

Db 1 VEEIAHQI 9

RESULT 12

AAE20865
 XX AAE20865 standard; peptide: 18 AA.

AC AAE20865;

XX 01-JUL-2002 (first entry)

XX E7-novispirin peptide.

XX Novispirin; antimicrobial; therapy; gram negative bacteria; infection;
 KW P. aeruginosa; S. maltophilia; Chlamydia trachomatis; ophthalmological;
 KW nonhaemolytic; tissue culture; cystic fibrosis; ophthalmic instillation;
 KW burn; bacterial vaginosis; sexually transmitted disease; antibacterial;
 KW plant-pathogenic pseudomonad; agricultural application; protozoacide;
 KW microbial infection; vulnery; fungicide.

XX Unidentified.

XX WO200200839-A2.

XX 03-JAN-2002.

XX 13-JUN-2001; 2001WO-US19094.

XX 28-JUN-2000; 2000US-0606858.

XX 19-APR-2001; 2001US-0840009.

XX (REGC) UNIV CALIFORNIA.

PA (UNIP) UNIV IOWA.

XX Lehrer RI, Waring AJ, Tack BF;

XX WPI; 2002-280458/32.

XX New antimicrobial novispirin peptides, useful for treating microbial
 PT infections caused by Gram-negative bacteria such as Pseudomonas
 PT aeruginosa, Chlamydia trachomatis, Escherichia coli or Stenotrophomonas
 PT maltophilia

XX Claim 4; Page 4; 42pp; English.

XX The present invention relates to an antimicrobial polypeptide (novispirin
 CC peptide) of a specific formula. Novispirin is useful for treating a
 CC microbial population comprising gram negative bacteria such as
 CC Pseudomonas aeruginosa, Chlamydia trachomatis, Escherichia coli or
 CC Stenotrophomonas maltophilia on contact with the microbial population.
 CC Novispirin mediated killing of microbes is also useful for modelling and
 CC screening novel antibiotics. Novispirin is useful for aerosol
 CC administration to lungs of patients with cystic fibrosis to treat
 CC infections caused by P. aeruginosa, S. maltophilia and to forestall the
 CC emergence of resistance to other inhaled antibiotics; instillation into
 CC the urinary bladder of patients with indwelling catheters to prevent
 CC infection; application to the skin of patients with serious burns;
 CC ophthalmic instillation, directly or in ophthalmic solutions, to treat or
 CC prevent infection; and intravaginal application to treat bacterial
 CC vaginosis and/or prevent sexually transmitted disease, e.g. by preventing
 CC infection with Chlamydia trachomatis. The novispirins also find use in
 CC the treatment of plant-pathogenic pseudomonads, in agricultural
 CC applications designed to prevent disease in and spoilage of food crops.
 CC Novispirin is useful in in vitro formulations to kill microbes, where the
 CC use of conventional antibiotics is not desirable, e.g., novispirins may
 CC be added to animal and/or human food preparations, and as an additive for
 CC in vitro cultures of cells, to prevent the overgrowth of microbes in
 CC tissue culture. Novispirin is also useful for killing non-bacterial
 CC pathogens such as fungal and protozoan pathogens. The invention is useful
 CC for treating a host suffering from or predisposed to a microbial
 CC infection. The peptides are nonhaemolytic, exhibit reduced in vitro
 CC cytotoxicity relative to other antimicrobial peptides and are well-
 CC tolerated in vivo after intravenous injection. Novispirins are equally
 CC effective against growing and stationary phase Pseudomonas aeruginosa and
 CC they retain activity in the presence of high concentrations of salt or

CC human serum. Novispirins also bind lipopolysaccharide (LPS), a property
 CC that may mitigate symptoms associated with gram-negative bacterial
 CC infection. The present sequence is E7-novispirin peptide.

XX Sequence 18 AA;

XX Query Match 30.8%; Score 32; DB 23; Length 18;

XX Best Local Similarity 50.0%; Pred. No. 2.5e+02;

XX Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 LEAKICHQIEEY 14

DB 6 IERKIHIIKKY 17

RESULT 13

AAE43395

ID AAE43395 standard; peptide; 8 AA.

XX AAE43395;

XX 12-MAY-1994 (first entry)

XX La/SSB epitope 24.

XX Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

XX Homo sapiens.

XX WO9321223-A.

XX 28-OCT-1993.

XX 13-APR-1993; 93WO-US03484.

XX 13-APR-1992; 92US-0867819.

XX (OKLA) UNIV OKLAHOMA STATE.

XX Harley JB;

XX WPI; 1993-351658/44.

XX New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus

XX Claim 1; Page 30; 43pp; English.

XX The sequences given in AAE43391-562 are linear epitopes which are
 CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polypyridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.

XX SQ Sequence 8 AA;
 Query Match 29.8%; Score 31; DB 14; Length 8;
 Best Local Similarity 100.0%; Pred. NO. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 YFGDF 19
 |||||
 Db 1 YFGDF 5

RESULT 14
 ABB84046
 ID ABB84046 standard; peptide; 10 AA.
 XX AC ABB84046;
 XX DT 21-AUG-2002 (first entry)
 XX DE Transglutaminase inhibitory peptide cr type #16.
 XX KW Transglutaminase inhibitor; Factor XIII inhibitor; XIIa inhibitor;
 KW ophthalmological; antiinflammatory; antirheumatic; antiarthritic;
 KW thrombolytic; neuroprotective; nootropic; antiseborrheic; dermatological;
 KW cytostatic; anti-HIV; antipsoriatic; cataract; inflammatory disease;
 KW arthritis; thrombosis; Alzheimer's disease; Huntington's chorea;
 KW acne; cancer; HIV infection; psoriasis.
 XX Unidentified.
 OS WO200236798-A2.
 PN 10-MAY-2002.
 XX 02-NOV-2001; 2001WO-EP12727.
 XX 03-NOV-2000; 2000DE-1054687.
 XX (NZYM-) N ZYME BIOTEC GMBH.
 XX Fuchsbaumer H, Pasternack R, Zotzel J;
 XX WPI; 2002-444364/47.
 XX New amino acid or peptide derivatives or analogs, are selective
 PT transglutaminase inhibitors useful e.g. for treating cataract,
 PT inflammatory diseases, rheumatoid arthritis, thrombosis, Alzheimer's
 PT disease and cancer
 XX Disclosure; Page 13; 44pp; German.

CC This invention describes novel amino acid or peptide derivatives or
 CC analogues (I), containing a modified side-chain (e.g. containing a formyl
 CC group) which are transglutaminase inhibitors and Factor XIII/XIIa
 CC inhibitors. The products of the invention have ophthalmological,
 CC antiinflammatory, antirheumatic, antithrombotic, thrombolytic,
 CC neuroprotective, nootropic, antiseborrheic, dermatological, cytostatic,
 CC anti-HIV and antipsoriatic activity. (I) are transglutaminase inhibitors,
 CC especially inhibitors of crosslinking of proteins or peptides
 CC (specifically fibrin and/or alpha₂-plasmin inhibitor), incorporation of
 CC primary amines in proteins and peptides, hydrolysis of the
 CC gamma-carboxamide group of glutamine residues bound in proteins or
 CC peptides, blood factor XIII/XIIa and mammalian, human, tissue, liver,
 CC brain, eye lens, keratinocyte, epidermal, prostate, plant, parasitic
 CC and/or bacterial transglutaminases. The products of the invention can be
 CC used for treating cataract, inflammatory diseases, rheumatoid arthritis,
 CC chronic arthritis, thrombosis; Alzheimer's disease, Huntington's chorea,
 CC acne, cancer (by induction of apoptosis), HIV infections and psoriasis.
 CC (I) Are targeted and specific transglutaminase inhibitors, which can
 CC inhibit a specific type of transglutaminase in the human or animal body
 CC without affecting other transglutaminases. ABB84001-ABB84049 represent
 CC transglutaminase inhibitors described in the method of the invention.

XX SQ Sequence 10 AA;
 Query Match 29.8%; Score 31; DB 23; Length 10;
 Best Local Similarity 71.4%; Pred. NO. 1.9e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 EAKICHQ 10
 :| ||||
 Db 3 QAPICHQ 9

RESULT 15
 AAE22202
 ID AAE22202 standard; peptide; 10 AA.
 XX AC AAE22202;
 XX DT 25-JUL-2002 (first entry)
 XX DE Murine MC-1 antibody heavy chain variable domain (VH) CDR3 peptide.
 XX KW Chemokine construct; human immunodeficiency virus 1; allergic disease;
 KW skin disease; immunological disorder; autoimmune disease; psoriasis;
 KW multiple sclerosis; rheumatoid arthritis; inflammatory bowel disease;
 KW diabetes; skin inflammation; atopic dermatitis; inflammatory disease;
 KW inflammatory renal disease; HIV-1; transplant rejection; murine; CDR3;
 KW heavy chain variable domain; VH; complementarity determining region 3;
 KW MC-1; antibody.
 XX Mus sp.
 OS WO200220615-A2.
 PN 14-MAR-2002.
 XX 10-SEP-2001; 2001WO-EP10433.
 XX 08-SEP-2000; 2000EP-0119694.
 XX 05-SEP-2001; 2001US-0948004.
 XX (MICR-) MICROMET AG.
 XX Mack M, Schloendorff D, Spring M;
 XX WPI; 2002-362240/39.
 XX Use of an antibody and/or chemokine construct that binds to a chemokine
 PT receptor, for eliminating cells latently infected with primate
 PT immunodeficiency virus, or treating, preventing and alleviating immune
 PT disorders
 XX Disclosure; Page 117; 117pp; English.

CC The invention relates to the use of an antibody and/or chemokine
 CC construct that binds a chemokine receptor for preparing a pharmaceutical
 CC composition for eliminating cells latently infected with a primate
 CC immunodeficiency virus (e.g. human immunodeficiency virus (HIV-1)). They
 CC are used in gene therapy and as vaccines. The antibody and/or chemokine
 CC construct is also used for preparing a pharmaceutical composition for
 CC treating, preventing and/or alleviating immunological disorders
 CC including autoimmune diseases (e.g. multiple sclerosis, type 1 diabetes
 CC and rheumatoid arthritis), allergic diseases, skin diseases (e.g. skin
 CC inflammation, atopic dermatitis and psoriasis), inflammatory diseases
 CC such as inflammatory joint diseases (chronic arthritis), inflammatory
 CC renal diseases and inflammatory bowel diseases and graft versus host
 CC and transplant rejections. The present sequence is murine MC-1 antibody
 CC heavy chain variable domain (VH) complementarity determining region 3
 CC (CDR3) peptide.
 XX Sequence 10 AA;
 SQ Query Match 29.8%; Score 31; DB 23; Length 10;

Best Local Similarity 71.48; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 13 EYFQDF 19
|||:|
Db 2 EYFQIF 8

Search completed: April 23, 2003, 13:42:59
Job time : 29.7135 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:21:27 ; Search time 27.9101 seconds
(without alignments)
132.885 Million cell updates/sec

Title: US-09-836-073-15
Perfect score: 100
Sequence: 1 LLDLTKICEIYVGFDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL-21.*
1: sp-archaea.*
2: sp-bacteria.*
3: sp-fungi.*
4: sp-human.*
5: sp-invertebrate.*
6: sp-mammal.*
7: sp-mhc.*
8: sp-organelle.*
9: sp-phage.*
10: sp-plant.*
11: sp-rodent.*
12: sp-virus.*
13: sp-vertebrate.*
14: sp-unclassified.*
15: sp-rvivirus.*
16: sp-bacteriaph.*
17: sp-archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	76.0	381	11 Q9CYB9	Q9CYB9 mus musculus
2	68	68.0	206	13 Q8QH15	Q8QH15 gallus gall
3	58	58.0	396	5 O01806	O01806 caenorhabdi
4	55	55.0	826	10 Q940X9	Q940X9 arabidopsis
5	54	54.0	391	2 P72382	P72382 staphylococ
6	54	54.0	391	2 P95709	P95709 staphylococ
7	54	54.0	391	16 Q99X57	Q99X57 staphylococ
8	53	53.0	411	10 Q9FL36	Q9FL36 arabidopsis
9	53	53.0	422	10 Q94A38	Q94A38 arabidopsis
10	53	53.0	1156	12 O57230	O57230 vaccinia vi
11	53	53.0	1164	12 Q90031	Q90031 vaccinia vi
12	53	53.0	1164	12 Q9JF79	Q9JF79 vaccinia vi
13	53	53.0	1164	12 Q90027	Q90027 vaccinia vi
14	53	53.0	1164	12 Q8V4V3	Q8V4V3 monkeypox v
15	53	53.0	1164	12 Q8V2N1	Q8V2N1 camelypox v
16	50	50.0	183	16 Q98Q96	Q98Q96 mycoplasma

17	48	48.0	190	12 Q90149	Q90149 bombyx mori
18	48	48.0	337	2 Q8RX98	Q8RX98 mycoplasma
19	48	48.0	390	5 Q8TRV5	Q8TRV5 drosophila
20	48	48.0	1222	12 Q92455	Q92455 bombyx mori
21	48	48.0	1222	12 P90691	P90691 bombyx mori
22	47	47.0	166	5 O44678	O44678 caenorhabdi
23	47	47.0	399	10 O8S0T8	O8S0T8 oryza sativ
24	47	47.0	545	10 O80567	O80567 arabidopsis
25	46	46.0	296	11 Q9CTN3	Q9CTN3 mus musculu
26	46	46.0	491	4 Q9BRS8	Q9BRS8 homo sapien
27	46	46.0	492	11 Q9D3J0	Q9D3J0 mus musculu
28	45	45.5	928	10 Q9LJ02	Q9LJ02 oryza sativ
29	45	45.0	343	10 Q9ALD0	Q9ALD0 oryza sativ
30	45	45.0	367	10 Q9SEU9	Q9SEU9 brassica na
31	45	45.0	399	10 O64531	O64531 arabidopsis
32	45	45.0	455	10 Q9LHL3	Q9LHL3 arabidopsis
33	45	45.0	1155	12 Q9Q8W4	Q9Q8W4 shope fibro
34	45	45.0	1155	12 Q9Q8I6	Q9Q8I6 myxoma viru
35	44	44.0	165	5 Q19281	Q19281 caenorhabdi
36	44	44.0	187	16 Q8XM19	Q8XM19 clostridium
37	44	44.0	343	2 Q9XB17	Q9XB17 bacillus ce
38	44	44.0	523	10 Q94K80	Q94K80 arabidopsis
39	44	44.0	1069	11 O88185	O88185 mus musculu
40	44	44.0	1743	5 Q9XW5	Q9XW5 caenorhabdi
41	43	43.0	132	16 Q926P7	Q926P7 listeria in
42	43	43.0	292	10 O80798	O80798 arabidopsis
43	43	43.0	465	5 Q9W3J4	Q9W3J4 drosophila
44	43	43.0	505	17 Q980D8	Q980D8 sulfolobus
45	43	43.0	658	10 Q9CAN3	Q9CAN3 arabidopsis

ALIGNMENTS

RESULT 1

Q9CYB9	PRELIMINARY;	PRT;	381 AA.
ID	Q9CYB9		
AC	Q9CYB9;		
DT	01-JUN-2001 (T-EMBLrel. 17, Created)		
DT	01-JUN-2001 (T-EMBLrel. 17, Last sequence update)		
DT	01-JUN-2002 (T-EMBLrel. 21, Last annotation update)		
DE	Sjogren syndrome antigen B.		
GN	SSB.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	SPRAIN=C57BL/6J; TISSUE=EMBRYO;		
RX	MDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,		
RA	Schirni L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustinch S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,		
RA	Havashizaki Y.		
RT	"Functional annotation of a full-length mouse cDNA collection."		
RL	Nature 409:685-690(2001).		
DR	EMBL; AK017822; BAB30957.1; -.		
DR	MGD; MGI:98423; SSB.		


```

DR InterPro: IPR002344; Lupus-La.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 1.
DR PRINTS; PR00302; LOPUSLA.
SQ SEQUENCE 381 AA; 43891 MW; 2E2DEF1452C0F0E9 CRC64;

Query Match 76.0%; Score 76; DB 11; Length 381;
Best Local Similarity 81.2%; Pred. No. 0.00051;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGDF 18
   : : : : : : : : : : : : : : : :
Db 13 LEAKICHOIEYFGDF 28

RESULT 2
Q8QH15 PRELIMINARY; PRT; 206 AA.
AC Q8QH15;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Untranslated region binding-protein.
GN UBP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA L'Ecuier T.J., Fang H.-L.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF467897; AAL76269.1; -
SQ SEQUENCE 206 AA; 23992 MW; 965B62F7DFFB90E9 CRC64;

Query Match 68.0%; Score 68; DB 13; Length 206;
Best Local Similarity 73.3%; Pred. No. 0.0053;
Matches 11; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGD 17
   : : : : : : : : : : : : : : :
Db 13 LESKICQIEYFGN 27

RESULT 3
O01806 PRELIMINARY; PRT; 396 AA.
AC O01806;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE C4E4.4 protein.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium. ";
RN Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RA Sammons L., Wohlmann P., Gillam B.;
RL "The sequence of C. elegans cosmid C4E4. ";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]

RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RL Waterston R.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003140; AAB54169.1; -
DR InterPro: IPR002344; Lupus-La.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 1.
DR PRINTS; PR00302; LOPUSLA.
DR SMART; PR00302; LOPUSLA.
DR SMART; SMO0360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS0030; RRM_RNP_1; FALSE_NEG.
SQ SEQUENCE 396 AA; 43631 MW; A2D828A4FAA3C34 CRC64;

Query Match 58.0%; Score 58; DB 5; Length 396;
Best Local Similarity 62.5%; Pred. No. 0.39;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LLDTKICEQIEYFGD 17
   : : : : : : : : : : : : : :
Db 11 DADQRIKQIEYFGN 26

RESULT 4
Q940X9 PRELIMINARY; PRT; 826 AA.
ID Q940X9
AC Q940X9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE A75921160/T10F18_190.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones. ";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY052365; AAK96556.1; -
DR InterPro: IPR001950; TIF_SUI1.
DR PROSITE; PS01118; SUI1_1; UNKNOWN_1.
SQ SEQUENCE 826 AA; 91377 MW; 8D41922E5B609D9A CRC64;

Query Match 55.0%; Score 55; DB 10; Length 826;
Best Local Similarity 52.9%; Pred. No. 2.4;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLDTKICEQIEYFGD 17
   : : : : : : : : : : : : : :
Db 276 LDLRDLKQIEYFGD 292

RESULT 5
P72382 PRELIMINARY; PRT; 391 AA.
ID P72382
AC P72382;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cap8P.
GN CAP8P.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

```

```

OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REYNOLDS, AND NEWMAN;
RX MEDLINE=96178981; PubMed=8606192;
RA Sau S., Lee C.Y.;
RT "Cloning of type 8 capsule genes and analysis of gene clusters for the
RT production of different capsular polysaccharides in Staphylococcus
RT aureus.";
RL J. Bacteriol. 178:2118-2126(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BECKER;
RX MEDLINE=97197525;
RA Sau S., Sun J., Lee C.Y.;
RT "Molecular characterization and transcriptional analysis of type 8
RT capsule genes in Staphylococcus aureus.";
RL J. Bacteriol. 179:1614-1621(1997).
DR EMBL; U73374; AAB49445.1; -
DR HSSP; P27828; 1F6D.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMs; TIGR00236; wecB; 1.
SQ SEQUENCE 391 AA; 44247 MW; 1E8D9FAA9BC76F0D CRC64;

Query Match 54.0%; Score 54; DB 2; Length 391;
Best Local Similarity 81.8%; Pred. No. 1.7;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 KICEQIEYYFG 16
DB 366 RICEAIEYYFG 376

RESULT 6
P95709
ID P95709 PRELIMINARY; PRT; 391 AA.
AC P95709;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-JUN-1997 (TRENBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Cap5P.
GN CAP5P.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REYNOLDS, AND NEWMAN;
RX MEDLINE=97388587; PubMed=9245821;
RA Sau S., Bhasin N., Wann E.R., Lee J.C., Foster T.J., Lee C.Y.;
RT "The Staphylococcus aureus allelic genetic loci for serotype 5 and 8
RT capsule expression contain the type-specific genes flanked by common
RT genes.";
RL Microbiology 143:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=REYNOLDS, AND NEWMAN;
RX MEDLINE=98101481; PubMed=9440531;
RA Kiser K.B., Lee J.C.;
RT "Staphylococcus aureus cap50 and cap5P genes functionally complement
RT mutations affecting enterobacterial common-antigen biosynthesis in
RT Escherichia coli.";
RL J. Bacteriol. 180:403-406(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=REYNOLDS, AND NEWMAN;
RX MEDLINE=98125727; PubMed=9466251;
RA Bhasin N., Albus A., Michon F., Livolsi P.J., Park J.-S., Lee J.C.;
RT "Identification of a gene essential for O-acetylation of the
RT Staphylococcus aureus type 5 capsular polysaccharide.";

```

```

RL Mol. Microbiol. 27:9-21(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=REYNOLDS, AND NEWMAN;
RA Bagga N., Wann E.R., Foster T.J., Lee J.C.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U81973; AAC46099.1; -
DR HSSP; P27828; 1F6D.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMs; TIGR00236; wecB; 1.
SQ SEQUENCE 391 AA; 44328 MW; B0105F690DB7CF1D CRC64;

Query Match 54.0%; Score 54; DB 2; Length 391;
Best Local Similarity 81.8%; Pred. No. 1.7;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 KICEQIEYYFG 16
DB 366 RICEAIEYYFG 376

RESULT 7
Q99X57
ID Q99X57 PRELIMINARY; PRT; 391 AA.
AC Q99X57;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Capsular polysaccharide synthesis enzyme Cap5P.
GN CAP OR SAV0164 OR SA0159.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Hiramatsu K.;
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AF003358; BAB56326.1; -
DR EMBL; AF003129; BAB41379.1; -
DR HSSP; P27828; 1F6D.
DR InterPro; IPR003331; Epimerase_2.
DR Pfam; PF02350; Epimerase_2; 1.
DR TIGRFAMs; TIGR00236; wecB; 1.
KW Complete proteome.
SQ SEQUENCE 391 AA; 44372 MW; D0DF5FA715BCCECC CRC64;

Query Match 54.0%; Score 54; DB 16; Length 391;
Best Local Similarity 81.8%; Pred. No. 1.7;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 6 KICEQIEYYFG 16
DB 366 RICEAIEYYFG 376

RESULT 8
Q9FL36
ID Q9FL36 PRELIMINARY; PRT; 411 AA.
AC Q9FL36;
DT 01-MAR-2001 (TRENBLrel. 16, Created)

```

```

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Similarity to RNA-binding protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98344145; PubMed=9679202;
RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT features of the regions of 1,381,565 bp covered by twenty one
RT physically assigned P1 and TAC clones."
RL DNA Res. 5:131-145(1998).
DR EMBL; AB010698; BAB11080.1; -
DR InterPro; IPR002344; Lupus_La.
DR IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR PRINTS; PR00302; LOPUSLA.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS50102; RRM; 1.
SQ SEQUENCE 411 AA; 45655 MW; A2EF62EB5B9B099 CRC64;

Query Match 53.0%; Score 53; DB 10; Length 411;
Best Local Similarity 56.2%; Pred. No. 2.5;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLDTKICEQIEYFGD 17
: : : : : : : : : : : : : : : :
Db 91 ELNQKIIRQVEYFSD 106

RESULT 9
Q9A38 PRELIMINARY; PRT; 422 AA.
AC Q9A38;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE A75946250/MPL12_3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY050403; AAK91419.1; -
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
DR PROSITE; PS50102; RRM; 1.
SQ SEQUENCE 422 AA; 46842 MW; 4EC4BBBF1E068F0E CRC64;

Query Match 53.0%; Score 53; DB 10; Length 422;
Best Local Similarity 56.2%; Pred. No. 2.6;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLDTKICEQIEYFGD 17
: : : : : : : : : : : : : : : :

```

```

Db 102 ELNQKIIRQVEYFSD 117

RESULT 10
Q57230 PRELIMINARY; PRT; 1156 AA.
AC Q57230;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE RNA polymerase subunit rpol32.
GN MVA135R.
OS Vaccinia virus (strain Ankara).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=126794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ANKARA;
RA Antoine G., Scheiflinger F., Falkner F.G., Dörner F.;
RT "The complete genomic sequence of the Modified Vaccinia Ankara (MVA)
RT strain";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE -> N DIPHOSPHATE +
CC [RNA](N).
DR EMBL; U94848; AAB96526.1; -
DR InterPro; IPR001572; RNA_pol_B.
DR Pfam; PF00562; RNA_pol_B; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
KW DNA-directed RNA polymerase; Transcription; Transferase.
SQ SEQUENCE 1156 AA; 132425 MW; F6657C8AF5E22EC3 CRC64;

Query Match 53.0%; Score 53; DB 12; Length 1156;
Best Local Similarity 58.8%; Pred. No. 6.9;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 DLDTKICEQIEYFGD 17
: : : : : : : : : : : : : : : :
Db 495 LDLEKKICEYIRSYKD 511

RESULT 11
Q90031 PRELIMINARY; PRT; 1164 AA.
AC Q90031;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE ORF15R.
GN A25R.
OS Variola virus, and
OS Variola minor virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10255, 53258;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Variola virus; STRAIN=GARCIA-1966;
RA Sichelkounov S.N., Totmenin A.V., Sosnovtsev S.V., Safronov P.F.,
RA Resenchuk S.M., Blinov V.M., Sandakhchiev L.S.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=variola minor virus; STRAIN=GARCIA-1966;
RA Sichelkounov S.N., Totmenin A.V., Gutorov V.V., Safronov P.F.,
RA Massung R.F., Loparev V.N., Knight J.C., Chizhikov V.E., Parsons J.M.,
RA Esposito J.J., Sosnovtsev S.;
RT "Analysis of the complete coding sequence of DNA of alastrim variola
RT minor virus strain Garcia-1966";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE -> N DIPHOSPHATE +
CC [RNA](N).
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.

```

```

DR EMBL: X76268; CAA53897.1; -
DR EMBL: Y16780; CAB54728.1; -
DR InterPro: IPR001572; RNA_pol_B.
DR Pfam: PF00562; RNA_pol_B; 1.
DR PROSITE: PS01166; RNA_POL_BETA; 1.
KW DNA-directed RNA polymerase; Transcription; Transferase.
SQ SEQUENCE 1164 AA; 133389 MW; 92984FECE35AABE2 CRC64;

Query Match          53.0%; Score 53; DB 12; Length 1164;
Best Local Similarity 58.8%; Pred. No. 6.9;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEYFGD 17
DB 503 LDLEKKICEYIRSYYKD 519
      |||: |||| | | |
      1  LDLDTKICEQIEYFGD 17
      503 LDLEKKICEYIRSYYKD 519

RESULT 12
ID Q9JF79 PRELIMINARY; PRT; 1164 AA.
AC Q9JF79;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE TA25R.
OS Vaccinia virus (strain Tian Tan).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10253;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TIAN TAN;
RA Jin Q., Hou Y.D., Cheng N.H., Yao E.M., Cheng S.X., Yang X.K.,
RA Jing D.Y., Yu W.H., Yuan J.S., Ma X.J.;
RT "Complete genomic sequence of vaccinia virus (Tian Tan strain).";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE -> N DIPHOSPHATE +
CC [RNA](N).
DR EMBL: AF095689; AAF34020.1; -
DR InterPro: IPR001572; RNA_pol_B.
DR Pfam: PF00562; RNA_pol_B; 1.
DR PROSITE: PS01166; RNA_POL_BETA; 1.
KW DNA-directed RNA polymerase; Transcription; Transferase.
SQ SEQUENCE 1164 AA; 133327 MW; 48EDC190C53C9914 CRC64;

Query Match          53.0%; Score 53; DB 12; Length 1164;
Best Local Similarity 58.8%; Pred. No. 6.9;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEYFGD 17
DB 503 LDLEKKICEYIRSYYKD 519
      |||: |||| | | |
      1  LDLDTKICEQIEYFGD 17
      503 LDLEKKICEYIRSYYKD 519

RESULT 13
ID Q90027 PRELIMINARY; PRT; 1164 AA.
AC Q90027;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE A25R.
GN A25R.
OS Variola major virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=12870;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BANGLADESH-1975;
RX MEDLINE=94088747; PubMed=8264798;
RA Massung R.F., Esposito J.J., Liu L., Qi J., Utterback T.R.,
RA Knight J.C., Aubin L., Yuran T.E., Parsons J.M., Loparev V.N.,

```

```

RA Seliwanov N.A., Cavallaro K.F., Kerlavage A.R., Mahy B.W.J.,
RA Venter C.J.;
RT "Potential virulence determinants in terminal regions of variola
RT smallpox virus genome.";
RL Nature 366:748-751(1993).
CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE -> N DIPHOSPHATE +
CC [RNA](N).
DR EMBL: L22579; AAA60876.1; -
DR InterPro: IPR001572; RNA_pol_B.
DR Pfam: PF00562; RNA_pol_B; 1.
DR PROSITE: PS01166; RNA_POL_BETA; 1.
KW DNA-directed RNA polymerase; Transcription; Transferase.
SQ SEQUENCE 1164 AA; 133401 MW; D28A83F6EDB8101B CRC64;

Query Match          53.0%; Score 53; DB 12; Length 1164;
Best Local Similarity 58.8%; Pred. No. 6.9;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEYFGD 17
DB 503 LDLEKKICEYIRSYYKD 519
      |||: |||| | | |
      1  LDLDTKICEQIEYFGD 17
      503 LDLEKKICEYIRSYYKD 519

RESULT 14
ID Q8V4V3 PRELIMINARY; PRT; 1164 AA.
AC Q8V4V3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE A25R.
GN A25R.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ZAIRE-96-I-16;
RX MEDLINE=21592287; PubMed=11734207;
RA Shchelkunov S.N., Totmenin A.V., Babkin I.V., Safronov P.F.,
RA Ryazankina O.I., Petrov N.A., Gutorov V.V., Uvarova E.A.,
RA Mikheev M.V., Sisler J.R., Esposito J.J., Jahrling P.B., Moss B.,
RA Sandakhchiev L.S.;
RT "Human monkeypox and smallpox viruses: genomic comparison.";
RL FEBS Lett. 509:66-70(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ZAIRE-96-I-16;
RA Shchelkunov S.N., Totmenin A.V., Safronov P.F., Gutorov V.V.,
RA Ryazankina O.I., Petrov N.A., Babkin I.V., Uvarova E.A.,
RA Esposito J.J., Moss B., Sisler J.R., Jahrling P.B., Sandakhchiev L.S.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF380138; AAL40593.1; -
DR InterPro: IPR001572; RNA_pol_B.
DR Pfam: PF00562; RNA_pol_B; 1.
DR PROSITE: PS01166; RNA_POL_BETA; UNKNOWN_1.
SQ SEQUENCE 1164 AA; 133434 MW; 9C2665752044B9BB CRC64;

Query Match          53.0%; Score 53; DB 12; Length 1164;
Best Local Similarity 58.8%; Pred. No. 6.9;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEYFGD 17
DB 503 LDLEKKICEYIRSYYKD 519
      |||: |||| | | |
      1  LDLDTKICEQIEYFGD 17
      503 LDLEKKICEYIRSYYKD 519

RESULT 15
ID Q8V2N1 PRELIMINARY; PRT; 1164 AA.
AC Q8V2N1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)

```

```

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DN RNA polymerase subunit RPO132 (CMP141R).
OS Camelpox virus (strain CP-1).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=28873;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-M-96;
RA Alfonso C.L., Tulman E.R., Lu Z., Zsak L., Zaitsev V.L.,
RA Kerenbikova U.Z., Sandybaev N.T., Kutish G.F., Rock D.L.;
RT "The genome of camelpox virus.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CMS;
RX PubMed=11907336;
RA Gubser C., Smith G.L.;
RT "The sequence of camelpox virus shows it is most closely related to
RT variola virus, the cause of smallpox.";
RL J. Gen. Virol. 83:855-872(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-CMS;
RA Gubser C., Smith G.L.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF438165; AAL73849.1; -
DR EMBL; AY009089; AAG37633.1; -
DR InterPro; IPR001572; RNA_pol_B.
DR Pfam; PF00562; RNA_pol_B; 1.
DR PROSITE; PS01166; RNA_POL_BETA; UNKNOWN_1.
SQ SEQUENCE 1164 AA; 133478 MW; 0E47088EC0FA64C8 CRC64;

Query Match          53.0%; Score 53; DB 12; Length 1164;
Best Local Similarity 58.8%; Pred. No. 6.9;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDLDTKICEIEYFGD 17
   ||| |||| | |
Db 503 LDLEKKICEIYRSYKD 519

Search completed: April 23, 2003, 13:32:54
Job time : 29.9101 secs

```

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:42 ; Search time 4.95506 Seconds
(without alignments)
150.669 Million cell updates/sec

Title: US-09-836-073-15

Perfect score: 100
Sequence: 1 LDLTKICEIYVFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	94.0	427	1 LAB_XENLA	P28048 xenopus lae
2	88	88.0	428	1 LAB_XENLA	P28048 xenopus lae
3	76	76.0	404	1 LA_BOVIN	P10881 bos taurus
4	76	76.0	408	1 LA_HUMAN	P05455 homo sapien
5	76	76.0	415	1 LA_MOUSE	P32067 mus musculu
6	76	76.0	415	1 LA_RAT	P38656 rattus norv
7	53	53.0	1164	1 RPO2_COWPX	P17474 cowpox viru
8	53	53.0	1164	1 RPO2_VACCV	P19798 vaccinia vi
9	53	53.0	1164	1 RPO2_VARV	P33811 variola vir
10	51	51.0	1221	1 V143_NPVAC	P23407 autographa
11	48	48.0	383	1 LA_AEDAL	Q26457 aedes albop
12	48	48.0	390	1 LA_DROME	P40796 drosophila
13	45	45.0	506	1 TCNO_PETCR	Q43033 petroselinu
14	44	44.0	490	1 IF74_HUMAN	O14879 homo sapien
15	44	44.0	1069	1 PCD7_HUMAN	O60245 homo sapien
16	43	43.0	366	1 DRPH_THEIN	P22823 thermocactin
17	43	43.0	711	1 ETE2_MYXVL	Q9q8k4 myxoma viru
18	43	43.0	711	1 ETE2_SFVKA	Q9q8y2 shope fibro
19	43	43.0	2184	1 POLG_ECOLIF	O91734 e genome po
20	42	42.0	270	1 YC87_METJA	O58683 methanococc
21	41	41.0	298	1 LAH1_SCHPO	P87058 schizosacch
22	41	41.0	438	1 I277_HUMAN	Q9nmr2 homo sapien
23	40	40.0	176	1 IPIR_HAEIN	Q24529 haemophilus
24	40	40.0	466	1 SRO9_YEAST	P25567 saccharomyc
25	40	40.0	552	1 FET4_YEAST	P40988 saccharomyc
26	40	40.0	608	1 ALB1_SALSA	P21848 salmo salar
27	40	40.0	608	1 ALB2_SALSA	Q03156 salmo salar
28	40	40.0	2911	1 FBN2_HUMAN	P35556 homo sapien
29	39.5	39.5	297	1 YX01_CAEEL	Q11108 caenorhabdi
30	39.5	39.5	967	1 SYL_PYRHO	O58698 pyrococcus
31	39	39.0	191	1 Y948_METJA	Q58358 methanococc
32	39	39.0	264	1 GRAK_HUMAN	P49863 homo sapien
33	39	39.0	305	1 HEM6_VIBCH	Q9kvk4 vibrio chol

RESULT 1

ID	LAB_XENLA	STANDARD	PRT	427 AA
AC	P28049;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen homolog B).			
DE	Lupus La protein homolog B (La ribonucleoprotein B) (La autoantigen homolog B).			
GN	LAB1.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Oocyte;			
RX	MEDLINE=93287095; PubMed=8510143;			
RA	Scherly D., Stutz F., Lin-Marq N., Clarkson S.G.;			
RT	"La proteins from Xenopus laevis. cDNA cloning and developmental expression."			
RL	J. Mol. Biol. 231:196-204(1993).			
CC	- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT POLYMERASE III TRANSCRIPTS (BY SIMILARITY).			
CC	- SUBCELLULAR LOCATION: Nuclear (Probable).			
CC	- DEVELOPMENTAL STAGE: BARELY DETECTABLE IN STAGE I/II OOCYTES, ACCUMULATE IN STAGE III/IV OOCYTES, THEN EXHIBIT A ROUGHLY CONSTANT STEADY STATE LEVEL IN MATURE OOCYTES, EGGS, AND EARLY EMBRYOS.			
CC	- PTM: PHOSPHORYLATED (PROBABLE).			
CC	- MISCELLANEOUS: THERE ARE TWO FORMS OF LA, LAA AND LAB, IN XENOPUS.			
CC	- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.			
CC	- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; X68818; CAA48716.1; -			
DR	PIR; S3817; S3817.			
DR	PIR; S3817; S3817.			
DR	InterPro; IPR002344; Lupus_La.			
DR	InterPro; IPR000504; RNA_rec_mot.			
DR	Pfam; PF00076; rrm; 1.			
DR	PRINTS; PR00302; LUPUSLA.			
DR	SMART; SM00360; RRM; 1.			
DR	PROSITE; PS50102; RRM; 1.			
DR	PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.			
KW	RNA-binding; Nuclear protein; Phosphorylation.			

34	39	39.0	451	1	GD1L_YEAST	P39958 saccharomyc
35	39	39.0	482	1	LBP_RABIT	P17454 oryctolagus
36	39	39.0	781	1	APE2_SULTO	Q974n6 sulfolobus
37	39	39.0	988	1	PINH_ARATH	Q9xgw1 arabidopsis
38	39	39.0	995	1	DBPA_YEAST	Q12389 saccharomyc
39	39	39.0	2283	1	DPOE_MOUSE	Q9wvf7 mus musculu
40	38.5	38.5	355	1	NOD1_RHIME	O52618 rhizobium m
41	38.5	38.5	692	1	FLHA_ECOLI	P19223 escherichia
42	38	38.0	98	1	CILG_ECOLI	P77618 escherichia
43	38	38.0	124	1	YBAJ_ECOLI	P37611 escherichia
44	38	38.0	177	1	IPYR_NEIMA	Q9jvg3 neisseria m
45	38	38.0	177	1	IPYR_NEIMB	Q9kg04 neisseria m

ALIGNMENTS

FT DOMAIN 110 202 RNA-BINDING (RRM).
 FT DOMAIN 315 331 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 427 AA; 48995 MW; 45F3146F93A355 CRC64;

Query Match 94.0%; Score 94; DB 1; Length 427;
 Best Local Similarity 100.0%; Pred. No. 5.4e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LDLTKICEIYYFGD 17
 |||||
 Db 10 LDLTKICEIYYFGD 26

RESULT 2

LAAXENLA STANDARD; PRT; 428 AA.
 AC P28048;
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lupus La protein homolog A (La ribonucleoprotein A) (La autoantigen homolog A).
 GN LAAL.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Oocyte;
 RX MEDLINE=93287095; PubMed=8510143;
 RA Scherly D., Stutz F., Lin-Marq N., Clarkson S.G.;
 RT "La proteins from Xenopus laevis. cDNA cloning and developmental expression";
 RL J. Mol. Biol. 231:196-204(1993).
 CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT POLYMERASE III TRANSCRIPTS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: BARELY DETECTABLE IN STAGE I/II OOCYTES, ACCUMULATE IN STAGE III/IV OOCYTES, THEN EXHIBIT A ROUGHLY CONSTANT STEADY STATE LEVEL IN MATURE OOCYTES, EGGS, AND EARLY EMBRYOS.
 CC -1- PM: PHOSPHORYLATED (PROBABLE).
 CC -1- MISCELLANEOUS: THERE ARE TWO FORMS OF LA, LAA AND LAB, IN XENOPUS.
 CC -1- SIMILARITY: HIGH, TO MAMMALIAN LA PROTEIN.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X68817; CAA48715.1;
 DR PIR; S28545; S28545.
 DR PIR; S33818; S33818.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rim; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 203 RNA-BINDING (RRM).
 FT DOMAIN 196 212 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 316 332 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 SQ SEQUENCE 428 AA; 48864 MW; AEB3A38B7D2E3EC3 CRC64;

Query Match 88.0%; Score 88; DB 1; Length 428;
 Best Local Similarity 94.1%; Pred. No. 5.3e-07;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LDLTKICEIYYFGD 17
 |||||
 Db 11 LSDTKICEIYYFGD 27

RESULT 3

LA_BOVIN STANDARD; PRT; 404 AA.
 AC P10881;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lupus La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
 GN SSB.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pituitary;
 RX MEDLINE=89202037; PubMed=2468131;
 RA Chan E.K.L., Sullivan K.F., Tan E.M.;
 RT "Ribonucleoprotein SS-B/La belongs to a protein family with consensus sequences for RNA-binding";
 RL Nucleic Acids Res. 17:2233-2244(1989).
 CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S, AND 7-2 RNAs.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- PM: PHOSPHORYLATED. THE PHOSPHORYLATION SITES ARE AT THE C-TERMINAL PART OF THE PROTEIN.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X13698; CAA31986.1;
 DR PIR; S03849; S03849.
 DR InterPro; IPR002344; Lupus_La.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rim; 1.
 DR PRINTS; PR00302; LUPUSLA.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 DR RNA-binding; Nuclear protein; Phosphorylation.
 FT DOMAIN 111 187 RNA-BINDING (RRM).
 SQ SEQUENCE 404 AA; 46534 MW; 4EE30B5C262AD6A1 CRC64;

Query Match 76.0%; Score 76; DB 1; Length 404;
 Best Local Similarity 81.2%; Pred. No. 4.8e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LDTKICEIYYFGD 18
 |||||
 Db 13 LEAKICHQIYYFGD 28

[illegible]


```
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L00993; AAA39415.1; -.
CC EMBL: Y07951; CAA69249.1; -.
CC MGD; MGI:98423; Ssb.
CC InterPro: IPR002344; Lupus.La.
CC InterPro: IPR000504; RNA_rec_mot.
CC Pfam: PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS0102; RRM; 1.
CC RNA-binding; Nuclear protein; Phosphorylation.
CC KW RNA-binding; Nuclear protein; Phosphorylation.
CC FT DOMAIN 111 187 RNA-BINDING (RRM).
CC SQ SEQUENCE 415 AA; 47756 MW; 2D7519769FDC933 CRC64;

Query Match 76.0%; Score 76; DB 1; Length 415;
Best Local Similarity 81.2%; Pred. No. 4.9e-05;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 3 LDTKICEIQEYFGDF 18
I: ||| |||||
Db 13 LEAKICHOIEYFGDF 28

RESULT 6
LA_RAT
ID LA_RAT STANDARD; PRT; 415 AA.
AC P38656;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lupus La protein homolog (la ribonucleoprotein) (La autoantigen
DE homolog).
DE SSB OR SS-B.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93246255; PubMed=7916708;
RA Senzel I., Troester H., Bartsch H., Schwemmler M., Igloi G.L.,
RA Bachmann M.;
RT "Isolation of rat cDNA clones coding for the autoantigen SS-B/La:
RT detection of species-specific variations.";
RL Gene 126:265-268(1993).
CC -1- FUNCTION: LA PROTEIN PLAYS A ROLE IN THE TRANSCRIPTION OF RNA
CC POLYMERASE III. IT IS MOST PROBABLY A TRANSCRIPTION TERMINATION
CC FACTOR. BINDS TO THE 3' TERMINI OF VIRTUALLY ALL NASCENT
CC POLYMERASE III TRANSCRIPTS. IT IS ASSOCIATED WITH PRECURSOR FORMS
CC OF RNA POLYMERASE III TRANSCRIPTS INCLUDING TRNA AND 4.5S, 5S, 7S,
CC AND 7-2 RNAs.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
CC EMBL: X67859; CAA48043.1; -.
CC PIR: JCI1494; JCI1494.
CC InterPro: IPR002344; Lupus.La.
CC InterPro: IPR000504; RNA_rec_mot.
CC Pfam: PF00076; rrm; 1.
CC PRINTS; PR00302; LUPUSLA.
CC SMART; SM00360; RRM; 1.
CC PROSITE; PS0102; RRM; 1.
CC RNA-binding; Nuclear protein; Phosphorylation.
CC KW RNA-binding; Nuclear protein; Phosphorylation.
CC FT DOMAIN 111 187 RNA-BINDING (RRM).
CC SQ SEQUENCE 415 AA; 47777 MW; 033FD9CC1E475F98 CRC64;

Query Match 76.0%; Score 76; DB 1; Length 415;
Best Local Similarity 81.2%; Pred. No. 4.9e-05;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 3 LDTKICEIQEYFGDF 18
I: ||| |||||
Db 13 LEAKICHOIEYFGDF 28

RESULT 7
RPO2_COMPX
ID RPO2_COMPX STANDARD; PRT; 1164 AA.
AC P17474; Q90025;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DNA-directed RNA polymerase 132 kDa polypeptide (EC 2.7.7.6).
GN RPO132.
OS Cowpox virus (CPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10243;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-Brighton red;
RX MEDLINE=89125698; PubMed=2915377;
RA Patel D.D., Pickup D.J.;
RT "The second-largest subunit of the poxvirus RNA polymerase is similar
RT to the corresponding subunits of procaryotic and eucaryotic RNA
RT polymerases."
RL J. Virol. 63:1076-1086(1989).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THIS ENZYME CONSISTS OF AT LEAST EIGHT SUBUNITS.
CC -1- ALTERNATIVE PRODUCTS: HAS TWO TRANSCRIPTIONAL START SITES; ONE IS
CC OPERATIVE AT LATE TIMES ONLY, AND THE OTHER IS OPERATIVE BOTH AT
CC EARLY AND LATE TIMES.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M26173; AAA42919.1; -.
CC PIR: A31879; RNVZCP.
CC InterPro: IPR001572; RNA_pol_B.
CC Pfam: PF00562; RNA_pol_B; 1.
CC PROSITE; PS01166; RNA_POL_BETA; 1.
CC Transferase; DNA-directed RNA polymerase; Transcription; Zinc;
CC Zinc-finger; Alternative initiation.
CC FT CHAIN 1 1164 DNA-DIRECTED RNA POLYMERASE 132 KDA
```

FT CHAIN 9 1164 POLYPEPTIDE; LATE FORM SPECIFIC.
FT INIT_MET 9 9 DNA-DIRECTED RNA POLYMERASE 132 KDA
FT ZN_FING 1087 1106 POLYPEPTIDE; EARLY AND LATE FORM.
SQ SEQUENCE 1164 AA; 133323 MW; 220DE6EF238DBF7 CRC64;
Query Match 53.0%; Score 53; DB 1; Length 1164;
Best Local Similarity 58.8%; Pred. No. 0.9;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 LDLDTKICEQIEYFGD 17
DB 503 LDLEKKICEYIRSYKD 519
RESULT 8
RPO2_VACCV STANDARD; PRT: 1164 AA.
AC P19798;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase 132 kDa polypeptide (EC 2.7.7.6).
GN RPO132 OR A24R.
OS Vaccinia virus (strain WR), and
OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus
OX NCBI_TaxID=10254, 10249;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WR;
RX MEDLINE=91082452; PubMed=1824607;
RA Amegadzie B.Y., Holmes M.H., Cole N.B., Jones E.V., Earl P.L.,
RA Moss B.;
RT Identification, sequence, and expression of the gene encoding the
RT second-largest subunit of the vaccinia virus DNA-dependent RNA
RT polymerase.";
RT Virology 179:517-563(1990).
RL Virology 180:88-98(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Copenhagen;
RX MEDLINE=91021027; PubMed=2219722;
RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RT The complete DNA sequence of vaccinia virus.";
RT Virology 179:247-266(1990).
RN [3]
RP COMPLETE GENOME.
RC STRAIN-Copenhagen;
RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RT Appendix to 'The complete DNA sequence of vaccinia virus'.";
RL Virology 179:517-563(1990).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +
CC (RNA)(N).
CC STRAIN-Copenhagen;
RA Goebel S.J., Johnson G.P., Perkins M.E., Davis S.W., Winslow J.P.,
RA Paoletti E.;
RT Appendix to 'The complete DNA sequence of vaccinia virus'.";
RL Virology 179:517-563(1990).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THIS ENZYME CONSISTS OF AT LEAST EIGHT SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; M37415; AAA72882.1; -.
CC EMBL; M35027; AAA48148.1; -.

DR PIR: H42519; RN28T.
DR PIR: A38517; RN2WA.
DR InterPro: IPR001572; RNA_pol_B.
DR Pfam: PF00562; RNA_pol_B; 1.
DR PROSITE: PS01166; RNA_POL_BETA; 1.
KW Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger.
FT ZN_FING 1087 1106 C4-TYPE (POTENTIAL).
SQ SEQUENCE 1164 AA; 133363 MW; B1F5C1484BA37D0D CRC64;
Query Match 53.0%; Score 53; DB 1; Length 1164;
Best Local Similarity 58.8%; Pred. No. 0.9;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 LDLDTKICEQIEYFGD 17
DB 503 LDLEKKICEYIRSYKD 519
RESULT 9
RPO2_VARV STANDARD; PRT: 1164 AA.
AC P33811;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE DNA-directed RNA polymerase 132 kDa polypeptide (EC 2.7.7.6).
GN RPO132 OR A24R.
OS Variola virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10255;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-India-1967 / Isolate Ind3;
RX MEDLINE=93202281; PubMed=8384129;
RA Shchelkunov S.N., Blinov V.M., Sandakhchiev L.S.;
RT "Genes of variola and vaccinia viruses necessary to overcome the host
RT protective mechanisms.";
RL FEBS Lett. 319:80-83(1993).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THIS ENZYME CONSISTS OF AT LEAST EIGHT SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; X69198; CAA49069.1; -.
DR PIR: G36850; G36850.
DR InterPro: IPR001572; RNA_pol_B.
DR Pfam: PF00562; RNA_pol_B; 1.
DR PROSITE: PS01166; RNA_POL_BETA; 1.
KW Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger.
FT ZN_FING 1087 1106 C4-TYPE (POTENTIAL).
SQ SEQUENCE 1164 AA; 133329 MW; 0452B84ED810CD53 CRC64;
Query Match 53.0%; Score 53; DB 1; Length 1164;
Best Local Similarity 58.8%; Pred. No. 0.9;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 LDLDTKICEQIEYFGD 17
DB 503 LDLEKKICEYIRSYKD 519

```

RESULT 10
ID V143_NPVAC STANDARD; PRT; 1221 AA.
AC P24307;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Helicase.
GN P143.
OS Autographa californica nuclear polyhedrosis virus (ACMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=46015;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HR3;
RX MEDLINE=91134998; PubMed=1994581;
RA Lu A., Carstens E.B.;
RT "Nucleotide sequence of a gene essential for viral DNA replication in
the baculovirus Autographa californica nuclear polyhedrosis virus.";
RN Virology 181:336-347(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C6;
RX MEDLINE=94303173; PubMed=8030224;
RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT "The complete DNA sequence of Autographa californica nuclear
polyhedrosis virus.";
RN Virology 202:586-605(1994).
CC -1- FUNCTION: ESSENTIAL FOR THE INITIATION OF VIRAL DNA REPLICATION,
IT MAY CONTRIBUTE TO OTHER FUNCTIONS SUCH AS CONTROLLING THE
SWITCH TO THE LATE PHASE AND LEADING TO THE INHIBITION OF HOST
PROTEIN SYNTHESIS. REQUIRED FOR LATE AND VERY LATE GENE
EXPRESSION.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; M57687; AAA67907.1; -
DR EMBL; L22858; AAA66725.1; -
DR PIR; A38499; HUNVAV.
KW Helicase; DNA replication; ATP-binding; DNA-binding; Nuclear protein;
Early protein.
FT DOMAIN 692 701 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT NP_BIND 917 924 ATP (BY SIMILARITY).
FT DNA_BIND 967 981 H-T-H MOTIF (BY SIMILARITY).
FT MUTAGEN 934 934 V->M; DEFECTIVE IN THE SYNTHESIS OF VIRAL
DNA, LATE PROTEIN SYNTHESIS AND THE
SHUTOFF OF HOST PROTEIN SYNTHESIS AT THE
NONPERMISSIVE TEMPERATURE (MUTANT TS8).
FT CONFLICT 126 126 F -> S (IN REF. 1).
FT CONFLICT 1149 1149 F -> L (IN REF. 1).
SQ SEQUENCE 1221 AA; 143213 MW; 090E199855882D1B CRC64;

Query Match 51.0%; Score 51; DB 1; Length 1221;
Best Local Similarity 61.5%; Pred. No. 2;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4. DTKICEQIYYFG 16
Db 440 DTKLCVSLGYFG 452
|||||: |||||

RESULT 11
LA_AEDAL STANDARD; PRT; 383 AA.
AC Q26457;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
OS Aedes albopictus (Forest day mosquito).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
OC Culicoidae; Aedes.
OX NCBI_TaxID=7160;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96135233; PubMed=8551578;
RA Pardigon N., Strauss J.H.;
RT "Mosquito homolog of the La autoantigen binds to Sindbis virus RNA.";
RN J. Virol. 70:1173-1181(1996).
CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTION TERMINATION BY RNA
POLYMERASE III. BINDS RNA AND DNA. BINDS TO THE 3' END OF THE
MINUS STRAND OF SINDBIS VIRUS RNA. THIS MAY BE SIGNIFICANT FOR
SINDBIS-VIRUS RNA REPLICATION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. PRIMARILY NUCLEAR,, BUT SIGNIFICANT
AMOUNTS ARE PRESENT IN THE CYTOPLASM.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; S80954; AAB35931.1; -
DR InterPro; IPR002344; Lupus.La.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; yrm; 1
DR PRINTS; PR00302; LOPUSLA.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW RNA-binding; Nuclear protein; DNA-binding.
FT DOMAIN 141 228
SQ SEQUENCE 383 AA; 44430 MW; 4E5CC8F21C40F452 CRC64;

Query Match 48.0%; Score 48; DB 1; Length 383;
Best Local Similarity 53.3%; Pred. No. 1.9;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 LDTKICEQIYYFGD 17
Db 43 LEASTIRQLEYFGD 57
|||||: |||||

RESULT 12
LA_DROME STANDARD; PRT; 390 AA.
AC P40796; Q24375; Q9VIN2;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE La protein homolog (La ribonucleoprotein) (La autoantigen homolog).
GN LA OR CG10922.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S; TISSUE-Ovary.
RX MEDLINE=94309632; PubMed=8035794;

```

RA Bai C., Li Z., Tolias P.P.;
 RT "Developmental characterization of a Drosophila RNA-binding protein
 homologous to the human systemic lupus erythematosus-associated
 La/SS-B autoantigen";
 RL Mol. Cell. Biol. 14:5123-5129(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94309661; PubMed=8035818;
 RA Yoo C.J., Wolin S.L.;
 RT "La proteins from Drosophila melanogaster and Saccharomyces
 cerevisiae: a yeast homolog of the La autoantigen is dispensable for
 growth";
 RL Mol. Cell. Biol. 14:5412-5424(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Abmayr A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Rubeis B., Deichner A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Neilson D.R., Neilson K.A., Nixon K., Nusskern D.R., Pacile J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2193(2000).
 CC -1- FUNCTION: MAY BE INVOLVED IN TRANSCRIPTION TERMINATION BY RNA
 POLYMERASE III. BINDS RNA AND DNA. BINDS TO PRECURSORS OF RNA
 POLYMERASE III TRANSCRIPTS. MAY PLAY A SPECIALIZED ROLE DURING FLY
 DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT EMBRYONIC, LARVAL,
 CC PUPAL, AND ADULT DEVELOPMENT. EXPRESSION THROUGHOUT THE EMBRYO IS
 CC FOLLOWED BY A RESTRICTED PATTERN OF MESODERMAL EXPRESSION THAT IS
 CC LATER CONFINED TO THE VISCERAL MESODERM, GONADS, GUT, AND SALIVARY
 CC GLANDS.
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -1- SIMILARITY: TO VERTEBRATE PROTEIN LA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL: U07652; AAA20518.1;
 CC EMBL: L32988; AAA21776.1;
 CC EMBL: AE003666; AAF53885.1;
 CC EMBL: FBgn0011638; La.
 CC InterPro: IPR002344; Lupus_La.
 CC Pfam: PF00076; rtm; 1.
 CC PRINTS: PR00302; LUPUSLA.
 CC SMART: SM00360; RRM; 1.
 CC PROSITE: PS0102; RRM; 1.
 CC PROSITE: PS0030; RRM_RNP_1; 1.
 CC RRNA-binding; Nuclear protein; DNA-binding.
 CC FT DOMAIN 149 234
 CC FT CONFLICT 169 169 A -> T (IN REF. 1).
 CC FT CONFLICT 182 183 KH -> NS (IN REF. 1).
 CC FT CONFLICT 283 283 A -> R (IN REF. 1).
 CC FT CONFLICT 329 329 K -> N (IN REF. 1).
 CC FT CONFLICT 329 329 K -> N (IN REF. 1).
 CC SEQUENCE 390 AA; 44884 MW; A8099288B90446A5 CRC64;
 SQ
 Query Match 48.0%; Score 48; DB 1; Length 390;
 Best Local Similarity 72.7%; Pred. No. 1.9;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 7 ICQIEVYFGD 17
 Db 54 IIRQVEYFGD 64
 | :|:|:|:|:|
 RESULT 13
 TCMO_PETCR STANDARD; PRT; 506 AA.
 ID TCMO_PETCR
 AC Q43033;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trans-cinnamate 4-monooxygenase (EC 1.14.13.11) (Cinnamic acid
 DE 4-hydroxylase) (CA4H) (C4H) (P450C4H) (Cytochrome P450 73).
 GN CYP73A10 OR CYP73.
 OS Petroselinum crispum (Parsley) (Petroselinum hortense).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Apiales; Apiaceae; Petroselinum.
 OX NCBI_TaxID=4043;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95320184; PubMed=7597051;
 RA Logemann E., Farniske M., Hahlbrock K.;
 RT "Modes of expression and common structural features of the complete
 RT phenylalanine ammonia-lyase gene family in parsley.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:5905-5909(1995).
 CC -1- FUNCTION: CONTROLS CARBON FLUX TO PIGMENTS ESSENTIAL FOR
 CC POLLINATION OR UV PROTECTION, TO NUMEROUS PHYTOALEXINS SYNTHESIZED
 CC BY PLANTS WHEN CHALLENGED BY PATHOGENS, AND TO LIGNINS.
 CC -1- CATALYTIC ACTIVITY: Trans-cinnamate + NADPH + O(2) -> 4-
 CC hydroxycinnamate + NADP(+) + H(2)O.
 CC -1- PATHWAY: Phenylpropanoid metabolism; second step.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL: L38898; AAC41660.1;
 CC InterPro: IPR001128; Cytochrome_P450.
 CC Pfam: PF00067; P450; 1.
 CC PRINTS: PR00385; P450.
 CC -----

```
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Heme; NADP.
FT BINDING 448 448 HEME (BY SIMILARITY).
SQ SEQUENCE 506 AA; 58047 MW; 32F00EE959D69CCF CRC64;

Query Match 45.0%; Score 45; DB 1; Length 506;
Best Local Similarity 42.9%; Pred. No. 7.8;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 5 TKICEQIEYIFGDF 18
Db 215 SRLAQSFYHFGDF 228
::: |||

RESULT 14
IFT4_HUMAN STANDARD; PRT; 490 AA.
AC O14879; Q99634; Q9BSK7;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon-induced protein with tetratricopeptide repeats 4 (IFIT-4)
DE (Interferon-induced 60 kDa protein) (IFI-60K) (ISG-60) (CIG49)
DE (Retinoic acid-induced gene G protein) (RIG-G).
GN IFIT4 OR IFI60.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Forebrain;
RX MEDLINE=98054347; PubMed=93911139;
RA Zhu H., Cong J.P., Shenk T.;
RT "Use of differential display analysis to assess the effect of human
RT cytomegalovirus infection on the accumulation of cellular RNAs;
RT Induction of Interferon-responsive RNAs";
RL Proc. Natl. Acad. Sci. U.S.A. 94:13985-13990(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99047533; PubMed=9828129;
RA de Veer M.J., Sim H., Whistock J.C., Devenish R.J., Ralph S.J.;
RT "IFI60/ISG60/IFIT4, a new member of the human IFI54/IFIT2 family of
RT interferon-stimulated genes.";
RL Genomics 54:267-277(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Yu M., Tong J., Mao M., Chen S., Chen Z.;
RT "RIG-G, a novel gene induced by ATRA in acute promyelocytic
RT leukemia cells, is a new member of the ISG family.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE IFIT FAMILY.
CC -1- SIMILARITY: CONTAINS 8 TPR REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF026939; AAB95160.1;
CC EMBL; AF083470; AAC63524.1;
CC EMBL; U52513; AAB40606.1;
CC EMBL; BC001383; AAH01383.1;
CC EMBL; BC004977; AAH04977.1;
CC Genew; HGNC:5411; IFIT4.

DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Heme; NADP.
FT BINDING 448 448 HEME (BY SIMILARITY).
SQ SEQUENCE 506 AA; 58047 MW; 32F00EE959D69CCF CRC64;

Query Match 45.0%; Score 45; DB 1; Length 506;
Best Local Similarity 42.9%; Pred. No. 7.8;
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 5 TKICEQIEYIFGDF 18
Db 215 SRLAQSFYHFGDF 228
::: |||

RESULT 14
IFT4_HUMAN STANDARD; PRT; 490 AA.
AC O14879; Q99634; Q9BSK7;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Interferon-induced protein with tetratricopeptide repeats 4 (IFIT-4)
DE (Interferon-induced 60 kDa protein) (IFI-60K) (ISG-60) (CIG49)
DE (Retinoic acid-induced gene G protein) (RIG-G).
GN IFIT4 OR IFI60.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Forebrain;
RX MEDLINE=98054347; PubMed=93911139;
RA Zhu H., Cong J.P., Shenk T.;
RT "Use of differential display analysis to assess the effect of human
RT cytomegalovirus infection on the accumulation of cellular RNAs;
RT Induction of Interferon-responsive RNAs";
RL Proc. Natl. Acad. Sci. U.S.A. 94:13985-13990(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99047533; PubMed=9828129;
RA de Veer M.J., Sim H., Whistock J.C., Devenish R.J., Ralph S.J.;
RT "IFI60/ISG60/IFIT4, a new member of the human IFI54/IFIT2 family of
RT interferon-stimulated genes.";
RL Genomics 54:267-277(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Yu M., Tong J., Mao M., Chen S., Chen Z.;
RT "RIG-G, a novel gene induced by ATRA in acute promyelocytic
RT leukemia cells, is a new member of the ISG family.";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE IFIT FAMILY.
CC -1- SIMILARITY: CONTAINS 8 TPR REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF026939; AAB95160.1;
CC EMBL; AF083470; AAC63524.1;
CC EMBL; U52513; AAB40606.1;
CC EMBL; BC001383; AAH01383.1;
CC EMBL; BC004977; AAH04977.1;
CC Genew; HGNC:5411; IFIT4.

DR MIN; 604650; -.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR; 5.
DR SMART; SM00028; TPR; 3.
KW Repeat; TPR repeat; Interferon induction.
FT REPEAT 51 84 TPR 1.
FT REPEAT 94 127 TPR 2.
FT REPEAT 136 169 TPR 3.
FT REPEAT 172 206 TPR 4.
FT REPEAT 207 240 TPR 5.
FT REPEAT 241 274 TPR 6.
FT REPEAT 415 448 TPR 7.
FT REPEAT 450 481 TPR 8.
FT CONFLICT 44 44 F -> S (IN REF. 4; AAH04977).
FT CONFLICT 359 359 Q -> QO (IN REF. 2).
FT CONFLICT 435 435 MISSING (IN REF. 2).
SQ SEQUENCE 490 AA; 55984 MW; B9F042D4DF7151D2 CRC64;

Query Match 44.0%; Score 44; DB 1; Length 490;
Best Local Similarity 41.2%; Pred. No. 11;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 DLDTKICEQIEYIFGDF 18
Db 33 DLEDRVNCQIEFLNTEF 49
||: |||: ||
||: |||: ||

RESULT 15
PCD7_HUMAN STANDARD; PRT; 1069 AA.
AC O60245; O60246; O60247;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protocadherin 7 precursor (Brain-heart protocadherin) (BH-Pcdh).
GN PCDH7 OR BHPCDH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98777460; PubMed=9615233;
RA Yoshida K., Yoshitomo-Nakagawa K., Seki N., Sasaki M., Sugano S.;
RT "Cloning, expression analysis, and chromosomal localization of
RT BH-protocadherin (PCDH7), a novel member of the cadherin
RT superfamily.";
RL Genomics 49:458-461(1998).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; A/BH-Pcdh-a (shown here), B/BH-
CC Pcdh-b and C/BH-Pcdh-c; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Expressed predominantly in brain and heart and
CC at lower levels in various other tissues.
CC -1- SIMILARITY: CONTAINS 7 CADHERIN DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB006755; BAA25194.1;
CC EMBL; AB006756; BAA25195.1;
CC EMBL; AB006757; BAA25196.1;
CC HSSP; P15116; 1NCJ.
CC Genew; HGNC:8659; PCDH7.
CC MIN; 602988; -.
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 11.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 7.
```

DR PROSITE; PS00232; CADHERIN_1; 6.
DR PROSITE; PS0268; CADHERIN_2; 7.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal; Alternative splicing.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 1069 PROTOCADHERIN 7.
FT DOMAIN 29 879 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 880 900 POTENTIAL.
FT DOMAIN 901 1069 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 29 143 CADHERIN 1.
FT DOMAIN 144 308 CADHERIN 2.
FT DOMAIN 309 415 CADHERIN 3.
FT DOMAIN 424 535 CADHERIN 4.
FT DOMAIN 536 639 CADHERIN 5.
FT DOMAIN 640 742 CADHERIN 6.
FT DOMAIN 745 862 CADHERIN 7.
FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 689 689 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 747 747 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 780 780 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 822 822 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 840 840 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 845 845 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPIC 1059 1069 MRLHPYITVFG -> VRCIPNFKYREG (IN ISOFORM B).
FT VARSPIC 265 311 MISSING (IN ISOFORM C).
FT VARSPIC 1059 1069 MRLHPYITVFG -> PFRVTFVSVQPDPHQGLQSCYD
SGLESETPSKSSSGPRGLCALPLPEDNYERTPDGSDVSR
PLPDVALTGKTRCEDVGHSDSCWMPVTSRPERKKSOPKL
STFMPVDERGSOEKLANGEAAIMGDRNRLNKKLTSSYET
FSAASFKEENANPEIDPLTKTGEYKPSPVNLTTRREVIL
(IN ISOFORM C).
SQ SEQUENCE 1069 AA; 116104 MW; 315349BE3414F66D CRC64;

Query Match 44.0%; Score 44; DB 1; Length 1069;
Best Local Similarity 60.0%; Pred. No. 25;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 2 DLDTKICEQIEYYFG 16
||| : |||||
DB 334 DLDVGNGQIEYYFG 348

Search completed: April 23, 2003, 13:28:08
Job time : 5.95506 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:03 ; Search time 10.1124 Seconds
(without alignments)
171.119 Million cell updates/sec

Title: US-09-836-073-15

Perfect score: 100
Sequence: 1 LDLTKICEQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	94.0	427	1 S33817	ribonucleoprotein
2	88	88.0	428	1 S33818	ribonucleoprotein
3	76	76.0	404	1 S03849	ribonucleoprotein
4	76	76.0	408	1 A31888	ribonucleoprotein
5	76	76.0	415	1 JCI494	ribonucleoprotein
6	58	58.0	396	2 T30953	hypothetical prote
7	54	54.0	391	2 H89777	capsular polysacch
8	53	53.0	1156	2 T37411	RNA polymerase sub
9	53	53.0	1164	1 RNVT8T	DNA-directed RNA p
10	53	53.0	1164	1 RNVTCP	DNA-directed RNA p
11	53	53.0	1164	2 T28566	DNA-directed RNA p
12	53	53.0	1164	2 F72166	A25R protein - var
13	53	53.0	1164	2 G36850	A24R protein - var
14	51	51.0	1221	1 HJNVAV	helicase (EC 3.6.1
15	50	50.0	183	2 F90570	hypothetical prote
16	48	48.0	190	2 A47569	p143 DNA helicase
17	48	48.0	390	2 A53773	La/SS-B homolog D-
18	48	48.0	390	2 A53781	ribonucleoprotein
19	48	48.0	1222	2 T41835	DNA helicase p143
20	47	47.0	166	2 T32701	hypothetical prote
21	47	47.0	529	2 T00677	hypothetical prote
22	45	45.0	399	2 T01035	hypothetical prote
23	45	45.0	506	2 T14907	trans-cinnamate 4-
24	44	44.0	165	2 T29244	hypothetical prote
25	44	44.0	1069	2 T00043	BH-protocadherin-a
26	44	44.0	1072	2 T00041	BH-protocadherin-p
27	44	44.0	1200	2 T00042	BH-protocadherin p
28	44	44.0	1743	2 T26859	hypothetical prote
29	43	43.0	292	2 T02349	hypothetical prote

ALIGNMENTS

RESULT 1

S33817
ribonucleoprotein La.B - African clawed frog
N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C:Species: Xenopus laevis (African clawed frog)
C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: S33817; S28544
R:Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.
J. Mol. Biol. 231, 196-204, 1993
A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expressio
A:Reference number: S33817; MUID:93287095; PMID:8510143
A:Accession: S33817
A:Molecule type: mRNA
A:Residues: 1-427 <SCH>
A:Cross-references: EMBL:X68818; NID:g64875; PIDN:CAA48716.1; PID:g64876
C:Comment: This protein associates with a variety of small RNA molecules, most of
ay act as a transcription termination factor.
C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
C:Keywords: phosphoprotein; RNA binding
F:111-177/Domain: ribonucleoprotein repeat homology <RRM>
F:112-117/Region: RNA-binding RNP2 motif
F:150-157/Region: RNA-binding RNP1 motif
F:227-427/Domain: phosphorylated #status predicted <PHY>

Query Match 94.0%; Score 94; DB 1; Length 427;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LDLTKICEQIEYFGD 17
|||||
DB 10 LDLTKICEQIEYFGD 26

RESULT 2

S33818
ribonucleoprotein La.A - African clawed frog
N:Alternate names: autoantigen SS-B/La; ribonucleoprotein SS-B
C:Species: Xenopus laevis (African clawed frog)
C:Date: 03-Feb-1994 #sequence_revision 26-May-1994 #text_change 22-Jun-1999
C:Accession: S33818; S28545
R:Scherly, D.; Stutz, F.; Lin-Marq, N.; Clarkson, S.G.
J. Mol. Biol. 231, 196-204, 1993
A:Title: La proteins from Xenopus laevis. cDNA cloning and developmental expressio
A:Reference number: S33817; MUID:93287095; PMID:8510143
A:Accession: S33818
A:Molecule type: mRNA
A:Residues: 1-428 <SCH>
A:Cross-references: EMBL:X68817; NID:g64873; PIDN:CAA48715.1; PID:g64874
C:Comment: This protein associates with a variety of small RNA molecules, most of
ay act as a transcription termination factor.
C:Superfamily: ribonucleoprotein La; ribonucleoprotein repeat homology
C:Keywords: phosphoprotein; RNA binding

R. Chambers, U.C.; Keene, J.D.
Proc. Natl. Acad. Sci. U.S.A. 82, 2115-2119, 1985

A:Title: The complete DNA sequence of vaccinia virus.
A:Reference number: A42531; MUID:91021027; PMID:2219722
A:Contents: annotation; possible protein-coding frames
A:Note: neither amino acid nor nucleotide sequence is given
C:Superfamily: DNA-directed RNA polymerase 132K polypeptide
C:Keywords: leucine zipper; nucleotidyltransferase; transcription; zinc finger
F:484-505/Region: leucine zipper motif
F:1087-1105/Region: zinc finger CCCC motif

Query Match 53.0%; Score 53; DB 1; Length 1164;
Best Local Similarity 58.8%; Pred. No. 3.2;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LDLDTKICEQIEYFGD 17
|||: |||||: |||
Db 503 LDLEKKICEYRSYK 519

RESULT 10
RVZCP

DNA-directed RNA polymerase (EC 2.7.7.6) 132K chain - cowpox virus (strain Brighton Red)
C:Species: cowpox virus
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 11-Jun-1999
C:Accession: A31879
R:Patel, D.D.; Pickup, D.J.
J. Virol. 63, 1076-1086, 1989

A:Title: The second-largest subunit of the poxvirus RNA polymerase is similar to the core
A:Reference number: A31879; MUID:89125698; PMID:2915377
A:Accession: A31879
A:Molecule type: DNA
A:Residues: 1-1164 <PAT>
A:Cross-references: GB:M26173; NID:g323393; PIDN:AAA42919.1; PID:g323394

C:Comment: This enzyme consists of at least seven subunits whose molecular weights are 1
C:Genetics:
A:Gene: rpol32
C:Superfamily: DNA-directed RNA polymerase 132K polypeptide
C:Keywords: nucleotidyltransferase; transcription; zinc finger

Query Match 53.0%; Score 53; DB 1; Length 1164;
Best Local Similarity 58.8%; Pred. No. 3.2;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LDLDTKICEQIEYFGD 17
|||: |||||: |||
Db 503 LDLEKKICEYRSYK 519

RESULT 11
T28566

DNA-directed RNA polymerase (EC 2.7.7.6) 133K chain - variola major virus
C:Species: variola major virus
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C:Accession: T28566
R:Massung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin
Nature 366, 748-751, 1993
A:Title: Potential virulence determinants in terminal regions of variola smallpox virus
A:Reference number: Z20488; MUID:94088747; PMID:8264798
A:Accession: T28566
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1164 <MAS>

A:Cross-references: EMBL:L22579; NID:9623595; PIDN:AAA60876.1; PID:g439046
A:Experimental source: strain Bangladesh-1975
C:Superfamily: DNA-directed RNA polymerase 132K polypeptide
C:Keywords: nucleotidyltransferase; transcription

Query Match 53.0%; Score 53; DB 2; Length 1164;
Best Local Similarity 58.8%; Pred. No. 3.2;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LDLDTKICEQIEYFGD 17
|||: |||||: |||
Db 503 LDLEKKICEYRSYK 519

RESULT 12
F72166

A25R protein - variola minor virus (strain Garcia-1966)
C:Species: variola minor virus
C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 20-Jun-2000
C:Accession: F72166
R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.F.; Massung, R.F.;
submitted to GenBank, March 1998
A:Description: Analysis of the complete coding sequence of DNA of alastrim variola
A:Reference number: A72150
A:Accession: F72166
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1164 <SHC>
A:Cross-references: GB:Y16780; NID:g5830555; PIDN:CAB54728.1; PID:g5830689

A:Experimental source: strain Garcia-1966
C:Genetics:
A:Gene: A25R
C:Superfamily: DNA-directed RNA polymerase 132K polypeptide

Query Match 53.0%; Score 53; DB 2; Length 1164;
Best Local Similarity 58.8%; Pred. No. 3.2;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LDLDTKICEQIEYFGD 17
|||: |||||: |||
Db 503 LDLEKKICEYRSYK 519

RESULT 13
G36850

A24R protein - variola virus (strain India-1967)
C:Species: variola virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-Mar-2001
C:Accession: G36850
R:Binov, V.M.
submitted to GenBank, November 1992
A:Reference number: A36859
A:Accession: G36850
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1164 <BLI>
A:Cross-references: GB:X69198; NID:g456758; PIDN:CAA49069.1; PID:g297307

C:Superfamily: DNA-directed RNA polymerase 132K polypeptide

Query Match 53.0%; Score 53; DB 2; Length 1164;
Best Local Similarity 58.8%; Pred. No. 3.2;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 LDLDTKICEQIEYFGD 17
|||: |||||: |||
Db 503 LDLEKKICEYRSYK 519

RESULT 14
HJNVAV

helicase (EC 3.6.1.-) - Autographa californica nuclear polyhedrosis virus
C:Species: Autographa californica nuclear polyhedrosis virus, ACMPNV
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 19-Jan-2001
C:Accession: A38499; H72861
R:Lu, A.; Carstens, E.B.
Virology 181, 336-347, 1991

A:Title: Nucleotide sequence of a gene essential for viral DNA replication in the
A:Reference number: A38499; MUID:91134998; PMID:1994581
A:Accession: A38499
A:Molecule type: DNA
A:Residues: 1-1221 <LUA>
A:Cross-references: EMBL:M57687

R:Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.
Virology 202, 586-605, 1994
A:Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis

A:Reference number: A72850; MUID:94303173; PMID:8030224
 A:Accession: H72861
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-125, 'F', 127-1148, 'F', 1150-1221 <AYR>
 A:Cross-references: GB:L22858; NID:g510708; PIDN:AAA66725.1; PID:g559164
 C:Genetics:
 A:Gene: Ac-helicase
 C:Superfamily: AcNVPV helicase
 C:Keywords: ATP; DNA binding; DNA repair; DNA replication; hydrolase; nucleotide binding
 F:917-924/Region: nucleotide-binding motif A (P-loop)

Query Match 51.0%; Score 51; DB 1; Length 1221;
 Best Local Similarity 61.5%; Pred. No. 7;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 4 DTKICEQIEYVFG 16
 ||||| : |||||
 Db 440 DTKLCVSLGYVFG 452

RESULT 15

F90570
 hypothetical protein MYPV_4700 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
 C:Species: Mycoplasma pulmonis
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
 C:Accession: F90570
 R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
 Nucleic Acids Res. 29, 2145-2153, 2001
 A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
 A:Reference number: A99512; MUID:21267165; PMID:11353084
 A:Accession: F90570
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-183 <KUR>
 A:Cross-references: GB:AL445566; PID:g14089884; PIDN:CAC13643.1; GSPDB:GN00153
 A:Experimental source: strain UAB CTIP
 C:Genetics:
 A:Gene: MYPV_4700
 A:Genetic code: SGC3
 C:Superfamily: inorganic pyrophosphatase

Query Match 50.0%; Score 50; DB 2; Length 183;
 Best Local Similarity 47.1%; Pred. No. 1.4;
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Oy 2 DLDTKICEQIEYVFGDF 18
 ||||| : ||||| : :
 Db 111 DLDSKWLEEIKYFSNY 127

Search completed: April 23, 2003, 13:34:38
 Job time : 11.1124 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:27:28 ; Search time 10.3146 Seconds
(without alignments)
139.834 Million cell updates/sec

Title: US-09-836-073-15

Perfect score: 100

Sequence: 1 LDLDTRKICEQIEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA.*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	18	9 US-09-836-073-15	Sequence 15, Appl
2	76	76.0	17	9 US-09-836-073-13	Sequence 13, Appl
3	76	76.0	18	9 US-09-836-073-1	Sequence 1, Appl
4	76	76.0	18	9 US-09-836-073-14	Sequence 14, Appl
5	76	76.0	460	9 US-10-102-806-695	Sequence 695, App
6	74	74.0	18	9 US-09-836-073-2	Sequence 2, Appl
7	72	72.0	18	9 US-09-836-073-4	Sequence 4, Appl
8	70	70.0	18	9 US-09-836-073-9	Sequence 9, Appl
9	68	68.0	18	9 US-09-836-073-11	Sequence 11, Appl
10	68	68.0	18	9 US-09-836-073-12	Sequence 12, Appl
11	67	67.0	18	9 US-09-836-073-10	Sequence 10, Appl
12	65.5	65.5	19	9 US-09-836-073-16	Sequence 16, Appl
13	65	65.0	18	9 US-09-836-073-3	Sequence 3, Appl
14	64	64.0	18	9 US-09-836-073-5	Sequence 5, Appl
15	64	64.0	38	9 US-09-843-676-25	Sequence 25, Appl
16	64	64.0	38	9 US-09-766-253-25	Sequence 25, Appl
17	64	64.0	38	9 US-09-438-486-25	Sequence 25, Appl
18	64	64.0	38	9 US-10-053-758-25	Sequence 25, Appl
19	64	64.0	38	9 US-10-054-295-25	Sequence 25, Appl

Sequence 25, Appl
Sequence 7, Appl
Sequence 8, Appl
Sequence 17, Appl
Sequence 19, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 24, Appl
Sequence 26, Appl
Sequence 26, Appl
Sequence 26, Appl
Sequence 26, Appl
Sequence 26, Appl
Sequence 6, Appl
Sequence 35, Appl
Sequence 222, Appl
Sequence 1409, Appl
Sequence 101, Appl
Sequence 1, Appl
Sequence 19, Appl
Sequence 16, Appl
Sequence 30, Appl

64 64.0 38 9 US-10-054-611-25
61 61.0 18 9 US-09-836-073-7
60 60.0 18 9 US-09-836-073-8
55 55.0 18 9 US-09-836-073-17
54 54.0 16 9 US-09-836-073-19
50.5 50.5 37 9 US-09-843-676-24
50.5 50.5 37 9 US-09-766-253-24
50.5 50.5 37 9 US-09-438-486-24
50.5 50.5 37 9 US-10-053-758-24
50.5 50.5 37 9 US-10-054-295-24
50.5 50.5 37 9 US-10-054-611-24
48 48.0 39 9 US-09-843-676-26
48 48.0 39 9 US-09-766-253-26
48 48.0 39 9 US-09-438-486-26
48 48.0 39 9 US-10-053-758-26
48 48.0 39 9 US-10-054-295-26
45 45.0 18 9 US-09-836-073-6
42.5 42.5 1597 9 US-09-832-292-35
41 41.0 80 9 US-10-002-344A-222
41 41.0 461 10 US-09-764-864-1409
40 40.0 287 9 US-10-108-605-101
40 40.0 745 10 US-09-881-852-1
40 40.0 755 10 US-09-881-852-19
40 40.0 972 10 US-09-924-154-16
39.5 39.5 449 9 US-10-004-551-30

ALIGNMENTS

RESULT 1
US-09-836-073-15
; Sequence 15, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; TYPE: PRT
; ORGANISM: Xenopus
US-09-836-073-15

Query Match 100.0%; Score 100; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LDLDTRKICEQIEYFGDF 18
| | | | | | | | | | | | | | | | | |
Db 1 LDLDTRKICEQIEYFGDF 18

RESULT 2
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 17
 ; TYPE: PRT
 ; ORGANISM: Mouse
 US-09-836-073-13

Query Match 76.0%; Score 76; DB 9; Length 17;
 Best Local Similarity 81.2%; Pred. No. 2.2e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LDTKICEQIEYFGDF 18
 I: ||| |||||
 Db 2 LEAKICHQIEYFGDF 17

RESULT 3

US-09-836-073-1
 ; Sequence 1, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baldya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 22002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-1

Query Match 76.0%; Score 76; DB 9; Length 18;
 Best Local Similarity 81.2%; Pred. No. 2.3e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LDTKICEQIEYFGDF 18
 I: ||| |||||
 Db 3 LEAKICHQIEYFGDF 18

RESULT 4

US-09-836-073-14
 ; Sequence 14, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baldya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 22002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Bovine
 US-09-836-073-14

Query Match 76.0%; Score 76; DB 9; Length 18;
 Best Local Similarity 81.2%; Pred. No. 2.3e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LDTKICEQIEYFGDF 18
 I: ||| |||||
 Db 3 LEAKICHQIEYFGDF 18

RESULT 5

US-10-102-806-695
 ; Sequence 695, Application US/10102806
 ; Publication No. US20030054421A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA103PIC1
 ; CURRENT APPLICATION NUMBER: US/10/102,806
 ; CURRENT FILING DATE: 2002-03-22
 ; PRIOR APPLICATION NUMBER: 09/925,298
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05881
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 846
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 695
 ; LENGTH: 460
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-102-806-695

Query Match 76.0%; Score 76; DB 9; Length 460;
 Best Local Similarity 81.2%; Pred. No. 0.00054;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LDTKICEQIEYFGDF 18
 I: ||| |||||
 Db 65 LEAKICHQIEYFGDF 80

RESULT 6

US-09-836-073-2
 ; Sequence 2, Application US/09836073
 ; Patent No. US20020173475A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dasgupta, Asim
 ; APPLICANT: Das, S.
 ; APPLICANT: Baldya, Narayan
 ; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
 ; FILE REFERENCE: 22002054822
 ; CURRENT APPLICATION NUMBER: US/09/836,073
 ; CURRENT FILING DATE: 2002-10-24
 ; PRIOR APPLICATION NUMBER: 09/316,630
 ; PRIOR FILING DATE: 1999-05-21
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 18
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-09-836-073-2

Query Match 74.0%; Score 74; DB 9; Length 18;
 Best Local Similarity 75.0%; Pred. No. 4.6e-05;
 Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LDTKICEQIEYFGDF 18
 I: ||| |||||
 Db 3 LEAKICHQIEYFGDF 18

RESULT 7

US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4

Query Match 72.0%; Score 72; DB 9; Length 18;
Best Local Similarity 80.0%; Pred. No. 9.3e-05;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DTKICEQIEYYFGDF 18
; : ||| ||||| |||||
DB 4 EAKICHQIEYYFGDF 18

RESULT 8

US-09-836-073-9
; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match 70.0%; Score 70; DB 9; Length 18;
Best Local Similarity 80.0%; Pred. No. 0.00019;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYYFGD 17
; : ||| ||||| |||||
DB 3 LEAKICHQIEYYFGD 17

RESULT 9

US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match 68.0%; Score 68; DB 9; Length 18;
Best Local Similarity 75.0%; Pred. No. 0.00038;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYYFGDF 18
; : ||| ||||| |||||
DB 3 LEAKICHQIEYYFGDF 18

RESULT 10

US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match 68.0%; Score 68; DB 9; Length 18;
Best Local Similarity 75.0%; Pred. No. 0.00038;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYYFGDF 18
; : ||| ||||| |||||
DB 3 LEAKICHQIEYYFGDF 18

RESULT 11

US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match 67.0%; Score 67; DB 9; Length 18;
Best Local Similarity 75.0%; Pred. No. 0.00054;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYYFGDF 18
I: ||| ||||| |||
Db 3 LEAKICHQIEYYQGF 18

RESULT 12

US-09-836-073-16

; Sequence 16, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 22002054822

; CURRENT APPLICATION NUMBER: US/09/836.073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16

; LENGTH: 19

; TYPE: PRT

; ORGANISM: Rat

US-09-836-073-16

Query Match 65.5%; Score 65.5; DB 9; Length 19;
Best Local Similarity 76.5%; Pred. No. 0.00097;
Matches 13; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 3 LDTKICEQI-EYYFGDF 18
I: ||| || |||||
Db 3 LEAKICHQIEYYFGDF 19

RESULT 13

US-09-836-073-3

; Sequence 3, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baidya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 22002054822

; CURRENT APPLICATION NUMBER: US/09/836.073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-3

Query Match 65.0%; Score 65; DB 9; Length 18;
Best Local Similarity 68.8%; Pred. No. 0.0011;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYYFGDF 18
I: ||| |||||
Db 3 LQAKICHQIQYYFGDF 18

RESULT 14

US-09-836-073-5

; Sequence 5, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 22002054822
; CURRENT APPLICATION NUMBER: US/09/836.073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-5

Query Match 64.0%; Score 64; DB 9; Length 18;
Best Local Similarity 73.3%; Pred. No. 0.0016;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DTKICEQIEYYFGDF 18
I: || |||||
Db 4 EOKOCHQIEYYFGDF 18

RESULT 15

US-09-843-676-25

; Sequence 25, Application US/09843676

; Patent No. US20020164786A1

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; APPLICANT: Lingner, Joachim

; APPLICANT: Nakamura, Toru

; APPLICANT: Chapman, Karen B.

; APPLICANT: Morin, Gregg B.

; APPLICANT: Harley, Calvin

; APPLICANT: Andrews, William H.

; TITLE OF INVENTION: NO. US20020164786A1el Telomerase

; NUMBER OF SEQUENCES: 225

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: United States of America

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/843.676

; FILING DATE: 26-Apr-2001

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/854.050

; FILING DATE: 09-MAY-1997

; APPLICATION NUMBER: US 08/846.017

; FILING DATE: 25-APR-1997

; APPLICATION NUMBER: US 08/844.419

; FILING DATE: 18-APR-1997

; APPLICATION NUMBER: US 08/724.643

; FILING DATE: 01-OCT-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Apple, Randolph T.

; REGISTRATION NUMBER: 36,429

; REFERENCE/DOCKET NUMBER: 015389-002930US

; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20020164786A1 Relevant
; TOPOLOGY: No. US20020164786A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-843-676-25

```

```

Query Match      64.0%; Score 64; DB 9; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.0032;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 7 ICQIEYIFGD 17
Db 1 ICQIEYIFGD 11

```

```

Search completed: April 23, 2003, 13:38:20
Job time : 10.3146 secs

```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:22:24 ; Search time 11.2247 Seconds
(without alignments)
47.183 Million cell updates/sec

Title: US-09-836-073-15
Perfect score: 100
Sequence: 1 LDLDTKICEIQEYFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pap.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	76.0	18	4	US-09-316-630-3
2	76	76.0	18	4	US-09-316-630-4
3	64	64.0	38	3	US-08-851-843A-25
4	64	64.0	38	4	US-08-974-549A-215
5	64	64.0	38	4	US-08-854-050-25
6	64	64.0	38	4	US-09-430-323-25
7	60	60.0	38	4	US-08-974-549A-214
8	50.5	50.5	37	3	US-08-851-843A-24
9	50.5	50.5	37	4	US-08-854-050-24
10	50.5	50.5	37	4	US-09-430-323-24
11	48	48.0	39	3	US-08-851-843A-26
12	48	48.0	39	4	US-08-974-549A-216
13	48	48.0	39	4	US-08-854-050-26
14	48	48.0	39	4	US-09-430-323-26
15	40	40.0	590	1	US-08-448-186A-9
16	40	40.0	754	2	US-08-941-262-1
17	40	40.0	755	2	US-08-941-262-3
18	40	40.0	775	2	US-08-966-388-4
19	40	40.0	775	3	US-09-188-403-4
20	40	40.0	775	4	US-09-188-404-4
21	40	40.0	775	4	US-09-281-259-4
22	40	40.0	1956	4	US-08-843-417-2
23	40	40.0	1956	4	US-08-843-417-10
24	39	39.0	83	3	US-08-851-843A-9
25	39	39.0	83	4	US-08-974-549A-191
26	39	39.0	83	4	US-08-854-050-9
27	39	39.0	83	4	US-09-430-323-9

28	39	39.0	85	3	US-08-851-843A-11	Sequence 11, Appl
29	39	39.0	85	4	US-08-974-549A-193	Sequence 193, App
30	39	39.0	85	4	US-08-854-050-11	Sequence 11, Appl
31	39	39.0	85	4	US-09-430-323-11	Sequence 11, Appl
32	39	39.0	94	4	US-08-974-549A-164	Sequence 164, App
33	39	39.0	456	1	US-08-205-719-4	Sequence 4, Appl
34	39	39.0	456	4	US-08-431-517F-6	Sequence 6, Appl
35	39	39.0	482	4	US-08-431-517F-5	Sequence 5, Appl
36	39	39.0	482	6	5245013-2	Patent No. 5245013
37	39	39.0	560	3	US-08-851-843A-6	Sequence 6, Appl
38	39	39.0	560	4	US-08-854-050-6	Sequence 6, Appl
39	39	39.0	560	4	US-09-430-323-6	Sequence 6, Appl
40	38.5	38.5	693	4	US-09-564-805-234	Sequence 234, App
41	38	38.0	38	3	US-08-851-843A-27	Sequence 27, Appl
42	38	38.0	38	4	US-08-974-549A-217	Sequence 217, App
43	38	38.0	38	4	US-08-854-050-27	Sequence 27, Appl
44	38	38.0	38	4	US-09-430-323-27	Sequence 27, Appl
45	38	38.0	505	4	US-09-627-216A-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3
Query Match 76.0%; Score 76; DB 4; Length 18;
Best Local Similarity 81.2%; Pred. No. 4.9e-06;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 3 LDLDTKICEIQEYFGDF 18
Db 3 LEAKICHQIEYFGDF 18
|: ||| ||||| |||||
RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316,630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
; OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 76.0%; Score 76; DB 4; Length 18;
Best Local Similarity 81.2%; Pred. No. 4.9e-06;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Caps 0;

QY 3 LDTKICQIEYYRGDF 18
I: ||| |||||
Db 3 LEAKICQIEYYRGDF 18

RESULT 3
US-08-851-843A-25
; Sequence 25, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200

; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-851-843A-25

Query Match 64.0%; Score 64; DB 3; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.00094;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 7 ICEQIEYYFGD 17
|||||
Db 1 ICEQIEYYFGD 11

RESULT 4
US-08-974-549A-215
; Sequence 215, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/915,503
; FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-0026100S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 215:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-215

Query Match 64.0%; Score 64; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.00094;
Matches 11; Conservative 0; Mismatches 0; Indels 0;

QY 7 ICEQIEYYFGD 17
Db 1 ICEQIEYYFGD 11

RESULT 5
US-08-854-050-25
Sequence 25, Application US/08854050
Patent No. 6261836
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6261836el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-0029300S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-854-050-25

Query Match 64.0%; Score 64; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.00094;
Matches 11; Conservative 0; Mismatches 0; Indels 0;

QY 7 ICEQIEYYFGD 17
Db 1 ICEQIEYYFGD 11

RESULT 6
US-09-430-323-25
Sequence 25, Application US/09430323
Patent No. 6309867
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6309867el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/430,323
FILING DATE: 29-Oct-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-430-323-25

Query Match 64.0%; Score 64; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.00094;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 ICEQIEYYFGD 17
| | | | | | | | | |
Db 1 ICEQIEYYFGD 11

RESULT 7
US-08-974-549A-214
; Sequence 214, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER-READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,549A
; FILING DATE: 19-NOV-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/911,312
; FILING DATE: 14-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,951

FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 214:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-974-549A-214

Query Match 60.0%; Score 60; DB 4; Length 38;
Best Local Similarity 83.3%; Pred. No. 0.0041;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICEQIEYYFGDF 18
| | | | | | | | | |
Db 1 ICHQXEYFGDF 12

RESULT 8
US-08-851-843A-24
; Sequence 24, Application US/08851843A
; Patent No. 6093809
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6093809el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER-READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/851,843A
; FILING DATE: 06-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION:

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/724,643
;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-0029300S
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 37 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; US-08-851-843A-24

Query Match 50.5%; Score 50.5; DB 3; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICEQIEYFGDF 18
|||iiiiii
Db 1 ICHQ-EYFGDF 11

RESULT 9
US-08-854-050-24
; Sequence 24, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643

;; FILING DATE: 01-OCT-1996
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Apple, Randolph T.
;; REGISTRATION NUMBER: 36,429
;; REFERENCE/DOCKET NUMBER: 015389-0029300S
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 24:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 37 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: not relevant
;; MOLECULE TYPE: peptide
;; US-08-854-050-24

Query Match 50.5%; Score 50.5; DB 4; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICEQIEYFGDF 18
|||iiiiii
Db 1 ICHQ-EYFGDF 11

RESULT 10
US-09-430-323-24
; Sequence 24, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-0029300S
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-430-323-24

Query Match 50.5%; Score 50.5; DB 4; Length 37;
Best Local Similarity 83.3%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 7 ICQIEYFGDF 18
II I I I I I I I I
Db 1 ICQI-EYFGDF 11

RESULT 11

US-08-851-843A-26
Sequence 26, Application US/08851843A
Patent No. 6093809
GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. 6093809el Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,843A
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-0029300S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids

TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-851-843A-26

Query Match 48.0%; Score 48; DB 3; Length 39;
Best Local Similarity 72.7%; Pred. No. 0.36;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICQIEYFGD 17
I I I I I I I I
Db 1 ILRQVEYFGD 11

RESULT 12

US-08-974-549A-216
Sequence 216, Application US/08974549A
Patent No. 6166178
GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: WO PCT/US97/17885
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002610US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 216:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-974-549A-216

Query Match 48.0%; Score 48; DB 4; Length 39;
Best Local Similarity 72.7%; Pred. No. 0.36;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 ICEQIEYYFGD 17
Db 1 ILRQVEYYFGD 11

RESULT 13
US-08-854-050-26
; Sequence 26, Application US/08854050
; Patent No. 6261836
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6261836el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-854-050-26

Query Match 48.0%; Score 48; DB 4; Length 39;
Best Local Similarity 72.7%; Pred. No. 0.36;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 ICEQIEYYFGD 17
Db 1 ILRQVEYYFGD 11

RESULT 14
US-09-430-323-26
; Sequence 26, Application US/09430323
; Patent No. 6309867
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: No. 6309867el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/430,323
; FILING DATE: 29-Oct-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
```


Job time : 12.2247 secs

; INFORMATION FOR SEQ ID NO: 26:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 39 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: not relevant
 ; MOLECULE TYPE: peptide
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
 US-09-430-323-26

Query Match 48.0%; Score 48; DB 4; Length 39;
 Best Local Similarity 72.7%; Pred. No. 0.36;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 ICEQIEYFQD 17
 | :|||
 Db 1 ILRQVEYFQD 11

RESULT 15
 US-08-448-196A-9
 ; Sequence 9, Application US/08448196A
 ; Patent No. 5780594
 ; GENERAL INFORMATION:
 ; APPLICANT: CARTER, DANIEL C.
 ; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
 ; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
 ; TITLE OF INVENTION: RELATED PROTEINS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NASA
 ; STREET: MARSHALL SPACE FLIGHT CENTER
 ; CITY: HUNTSVILLE
 ; STATE: ALABAMA
 ; COUNTRY: USA
 ; ZIP: 35812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/448,196A
 ; FILING DATE: 23-MAY-1995
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BROAD JR., ROBERT L.
 ; REGISTRATION NUMBER: 18,757
 ; REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 205-544-0021
 ; TELEFAX: 205-544-0258
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 590 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE: N-terminal
 US-08-448-196A-9

Query Match 40.0%; Score 40; DB 1; Length 590;
 Best Local Similarity 46.2%; Pred. No. 1.4e+02;
 Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 DLDTKICEQIEY 14
 | :||| :||
 Db 380 DIDKITDIDY 392

Search completed: April 23, 2003, 13:36:33

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:15:23 ; Search time 28.5169 Seconds
(without alignments)
84.109 Million cell updates/sec

Title: us-09-836-073-15

Perfect score: 100

Sequence: 1 LLDLTKICEQIEYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	76.0	18	AA52200	Human la autoantigen
2	76	76.0	21	AAG01351	Human secreted pro
3	76	76.0	408	AAW03716	Human autoantigen
4	76	76.0	439	ABG08417	Novel human diagno
5	76	76.0	460	AA58987	Breast and ovarian
6	76	76.0	460	ABP41511	Human ovarian anti
7	65	65.0	21	AA43394	La/SSB epitope 17.
8	48	48.0	390	ABG65316	Drosophila melanog
9	47	47.0	913	AA47714	Arabidopsis thalia
10	47	47.0	923	AA47713	Arabidopsis thalia

11	47	47.0	993	21	AA47712	Arabidopsis thalia
12	46.5	46.5	412	23	AAU96342	Der HMW-map polype
13	46.5	46.5	462	23	AAU96341	Der HMW-map polype
14	46	46.0	492	21	AA43110	Human ORFX ORF2874
15	45	45.0	223	21	AAG17695	Arabidopsis thalia
16	45	45.0	223	21	AA46046	Arabidopsis thalia
17	45	45.0	243	21	AAG17694	Arabidopsis thalia
18	45	45.0	243	21	AA46045	Arabidopsis thalia
19	45	45.0	399	21	AA46044	Arabidopsis thalia
20	44	44.0	490	20	AAU05375	Human HCMV inducib
21	44	44.0	491	22	AAU33137	Novel human secret
22	43	43.0	231	21	AA437339	Arabidopsis thalia
23	43	43.0	256	21	AA437338	Arabidopsis thalia
24	43	43.0	292	21	AA437337	Arabidopsis thalia
25	43	43.0	292	21	AA437336	Arabidopsis thalia
26	43	43.0	330	21	AA437335	Arabidopsis thalia
27	43	43.0	465	22	AB858885	Drosophila melanog
28	42	42.0	1998	22	AB866040	Drosophila melanog
29	41.5	41.5	1379	22	AB868940	Drosophila melanog
30	41	41.0	301	21	AA43232	Human ORFX ORF2996
31	41	41.0	328	23	AB449912	Listeria monocytog
32	41	41.0	332	22	AB898826	Zinc finger protei
33	41	41.0	450	22	AB850212	Human transcriptio
34	41	41.0	459	22	AB11681	Human Zn finger pr
35	41	41.0	461	22	AAU16456	Human novel secret
36	41	41.0	528	23	ABP41616	Human ovarian anti
37	41	41.0	688	22	AB861544	Drosophila melanog
38	41	41.0	782	20	AA42773	Rat neuronal lmed
39	41	41.0	793	22	AA439996	Human polypeptide
40	41	41.0	827	22	AA41782	Human polypeptide
41	41	41.0	3132	22	AA51367	Japanese yam mosai
42	40.5	40.5	449	20	AAW98132	Heart protein kina
43	40	40.0	111	22	AAE01505	Human gene 22 enco
44	40	40.0	111	23	ABG63908	Human albumin fusi
45	40	40.0	158	21	AA455318	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AA52200

ID: AA52200 standard; peptide; 18 AA.

XX AA52200;

XX 14-MAR-2000 (first entry)

XX Human la autoantigen peptide (LAP).

XX La autoantigen; LAP; Internal ribosome entry site; IRES; translation;
 KW viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
 KW coronavirus; hepatitis virus; rhinovirus; adenovirus; coxsackie virus;
 KW parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
 KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
 KW vesicular stomatitis virus.

XX Homo sapiens.

XX WO9961613-A2.

XX 02-DEC-1999.

XX 21-MAY-1999; 99WO-US11281.

XX 22-MAY-1998; 98US-0086527.

XX (REGC) UNIV CALIFORNIA.

XX Das S, Dasgupta A;

XX WPI; 2000-062712/05.

PT New yeast inhibitory peptide useful for inhibiting viral protein
 translation and replication -
 XX
 PS Claim 5; Page 57; 81pp; English.
 CC This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (RNA see AA45200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rabdovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.
 XX
 SQ Sequence 18 AA;

Query Match 76.0%; Score 76; DB 21; Length 18;
 Best Local Similarity 81.2%; Pred. No. 3.8e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGDF 18
 I: ||| |||||
 Db 3 LEAKICHQIEYFGDF 18

RESULT 2
 AAG01351
 ID AAG01351 standard; Protein; 92 AA.
 XX
 AC AAG01351;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 5432.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GIST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 PS WPI; 2000-500381/45.
 DR N-PSDB; AAG01357.
 XX
 CC New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PT
 XX Claim 13; SEQ ID 5432; 71pp + CD-ROM; English.
 PS
 SS The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30

CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5',
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SQ Sequence 92 AA;

Query Match 76.0%; Score 76; DB 21; Length 92;
 Best Local Similarity 81.2%; Pred. No. 0.00023;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGDF 18
 I: ||| |||||
 Db 13 LEAKICHQIEYFGDF 28

RESULT 3
 AAW03716
 ID AAW03716 standard; protein; 408 AA.
 XX
 AC AAW03716;

XX 12-MAR-1997 (first entry)
 XX
 DE Human autoantigen La(SS-B).
 XX
 KW Autoimmune disease; La autoantigen; Sjogren's syndrome;
 KW systemic lupus erythematosus; diagnosis.
 XX
 OS Homo sapiens.
 XX
 PN US5541291-A.
 XX
 PD 30-JUL-1996.
 XX
 PF 31-DEC-1984; 84US-0687908.
 XX
 PR 27-MAY-1987; 87US-0054871.
 PR 31-DEC-1984; 84US-0687908.
 XX
 PA (UYDU-) UNIV DUKE.
 XX
 PI Keene JD;

XX WPI; 1996-362015/36.
 XX
 PT Auto-antigen U2-RNP, associated with severe polymyositis-scleroderma
 PT overlap syndrome - useful for diagnosis and treatment of autoimmune
 PT diseases
 XX
 PS Disclosure; Columns 15-16; 21pp; English.
 XX
 CC The human lupus antigen (La) is diagnostic for Sjogren's syndrome,
 CC as well as occurring in systemic lupus erythematosus patients. The
 CC La protein is clinically related to the Ro protein that is highly
 CC common among autoimmune patients. La and Ro antigens sometimes
 CC reside on the same cellular ribonucleoprotein particle; most La
 CC patients contain some Ro antibodies and vice versa. La cDNA has
 CC been isolated from a human liver library.
 XX
 SQ Sequence 408 AA;

Query Match 76.0%; Score 76; DB 17; Length 408;
 Best Local Similarity 81.2%; Pred. No. 0.0012;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGDF 18

Db 13 LEAKICQIEYVFGDF 28

RESULT 4
ABG08417
ID ABG08417 standard; Protein; 439 AA.

XX AC ABG08417;
XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #8408.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS72604.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

XX PS Claim 20; SEQ ID NO 38776; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 439 AA;

Query Match 76.0%; Score 76; DB 22; Length 439;

Best Local Similarity 81.2%; Pred. No. 0.0013;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYVFGDF 18

Db 43 LEAKICQIEYVFGDF 58

RESULT 5

AAB58987
ID AAB58987 standard; Protein; 460 AA.

XX AC AAB58987;

XX DT 27-MAR-2001 (first entry)

XX DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 695.

XX KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neutrophic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.

XX OS Homo sapiens.

XX PN WO200055173-A1.

XX PD 21-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US05881.

XX PR 12-MAR-1999; 99US-0124270.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX DR WPI; 2000-611515/58.

XX DR N-PSDB; AAF21890.

XX PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases

XX PS Claim 11; Page 1149-11150; 1299pp; English.

XX CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC neutrophic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antitumor; vulnary; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiac activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemia; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.

XX Sequence 460 AA;

Query Match 76.0%; Score 76; DB 21; Length 460;

Best Local Similarity 81.2%; Pred. No. 0.0014;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYVFGDF 18

Db 1; ||| |||||

Db 65 LEAKICHEIYFEGDF 80

RESULT 6

ID ABP41511 standard; Protein; 460 AA.

XX AC ABP41511;

XX DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HVVAF56, SEQ ID NO:2643.

XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

XX KW ovarian cancer; breast cancer; tumour; reproductive system disorder;

XX KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

XX KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;

XX KW inflammatory condition; immune disorder; blood disorder;

XX KW cardiovascular disorder; respiratory disorder; neurological disorder;

XX KW gastrointestinal disorder; urinary system disorder; drug screening;

XX KW gene therapy; chromosome mapping; forensic analysis;

XX KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;

XX KW antiinflammatory; gynaecological; reproductive.

XX OS Homo sapiens.

XX PN WO200200677-A1.

XX PD 03-JAN-2002.

XX PF 07-JUN-2001; 2001WO-US18569.

XX PR 07-JUN-2000; 2000US-209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX DR WPI; 2002-147878/19.

XX DR N-PSDB; ABQ54588.

XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,

XX PT useful in the prevention, treatment and diagnosis of cancer (e.g.

XX PT ovarian cancer); immune disorders, cardiovascular disorders and

XX PT neurological diseases -

XX PS Claim 11; SEQ ID NO 2643; 2922pp; English.

XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-

XX CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also

XX CC encompasses polypeptides 90% identical and polynucleotides 95% identical

XX CC to the sequences of the invention. The invention additionally relates to

XX CC recombinant vectors and host cells comprising human ovarian antigen

XX CC polynucleotides, antibodies against human ovarian antigens, and the use

XX CC of ovarian antigen polynucleotides and polypeptides in diagnosing,

XX CC treating, prognosing or preventing various ovary and/or breast-related

XX CC disorders. Such conditions include ovarian cancer and breast cancer, and

XX CC metastatic tumours of ovarian or breast origin, reproductive system

XX CC disorders (e.g., infertility, disorders of pregnancy, anovulation,

XX CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine

XX CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic

XX CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and

XX CC vaginitis), immune disorders (e.g., congenital and acquired

XX CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),

XX CC blood-related disorders (e.g., anaemia), cardiovascular disorders,

XX CC respiratory disorders, neurological disorders, gastrointestinal disorders

XX CC and urinary system disorders. Ovarian antigen polypeptides and

XX CC polynucleotides may also be used in screening for compounds which

XX CC modulate ovarian antigen expression or activity. The polynucleotides may

XX CC further be used for gene therapy, chromosome mapping, in the

XX CC identification of individuals and in forensic analysis, and the

XX CC polypeptides may be used as food additives or to prepare antibodies

XX CC useful in disease diagnosis, drug targeting and phenotyping. The present

CC sequence represents a human ovarian antigen of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 460 AA;

Query Match 76.0%; Score 76; DB 23; Length 460;

Best Local Similarity 81.2%; Pred. No. 0.0014;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEIYFEGDF 18

Db 65 LEAKICHEIYFEGDF 80

RESULT 7

AAAR43394

ID AAR43394 standard; peptide; 21 AA.

XX AC AAR43394;

XX DT 12-MAY-1994 (first entry)

XX DE La/SSB epitope 17.

XX KW Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;

XX KW nuclear ribonucleoprotein; rRNP; Sm B/B'; polypeptide; antigen; D;

XX KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B';

XX KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

XX OS Homo sapiens.

XX PN WO9321223-A.

XX PD 28-OCT-1993.

XX PF 13-APR-1993; 93WO-US03484.

XX PR 13-APR-1992; 92US-0867819.

XX PA (OKLA) UNIV OKLAHOMA STATE.

XX PI Harley JB;

XX DR WPI; 1993-351658/44.

XX PT New linear epitope(s) for human auto-antibodies - from the

XX PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used

XX PT for diagnosing and treating auto-immune disorders e.g. systemic

XX PT lupus erythematosus

XX PS Claim 1; Page 30; 43pp; English.

XX CC The sequences given in AAR43391-562 are linear epitopes which are

XX CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,

XX CC the 70 kD nuclear ribonucleoprotein (rRNP) and the Sm B/B'

XX CC polypeptide. These antigens are common in systemic lupus

XX CC erythematosus (SLE) and closely related disorders. The Ro/SSA family

XX CC of proteins has been shown to have several molecular forms which are

XX CC defined by the molecular weight of the antigen identified. The major

XX CC form has a molecular weight of 60 kD and two additional forms have

XX CC molecular weights of 52 and 54 kD. La/SSB is also a member of this

XX CC group of autoantibodies and binds small RNAs with a polyuridine

XX CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin

XX CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric

XX CC phosphoprotein which associates with RNA polymerase III transcripts.

XX CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and

XX CC U5 RNA. Anti-Sm antibodies may be directed against one or a

XX CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),

XX CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be

XX CC used for preventing, treating or screening autoimmune disorders,

XX CC especially SLE or Sjogrens syndrome (SS). They bind to a human

PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140354.
 PR 28-JUN-1999; 99US-0140695.
 PR 29-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 26-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 47.0%; Score 47; DB 21; Length 913;
 Best Local Similarity 66.7%; Pred. No. 98;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 KICEQIEYFGD 17
 |||:|||||
 Db 116 KIVNQVEYFSD 127

RESULT 10
 AAG47713
 ID AAG47713 standard; Protein; 923 AA.
 XX
 AC AAG47713;
 XX
 DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 60167.
DE Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.

PR	27-AUG-1999;	99US-0151080.	OS	Arabidopsis thaliana.
PR	30-AUG-1999;	99US-0151303.	XX	
PR	31-AUG-1999;	99US-0151438.	PN	EP1033405-A2.
PR	01-SEP-1999;	99US-0151930.	XX	
PR	07-SEP-1999;	99US-0152363.	PD	06-SEP-2000.
PR	10-SEP-1999;	99US-0153070.	XX	
PR	13-SEP-1999;	99US-0153758.	PF	25-FEB-2000; 2000EP-0301439.
PR	15-SEP-1999;	99US-0154018.	XX	
PR	16-SEP-1999;	99US-0154039.	XX	99US-0121825.
PR	20-SEP-1999;	99US-0154779.	PR	99US-0123180.
PR	22-SEP-1999;	99US-0155139.	PR	99US-0123548.
PR	23-SEP-1999;	99US-0155486.	PR	99US-0125788.
PR	24-SEP-1999;	99US-0155659.	PR	99US-0126264.
PR	28-SEP-1999;	99US-0156458.	PR	99US-0126785.
PR	29-SEP-1999;	99US-0156596.	PR	99US-0127462.
PR	04-OCT-1999;	99US-0157117.	PR	99US-0128234.
PR	08-OCT-1999;	99US-0157753.	PR	99US-0128714.
PR	08-OCT-1999;	99US-0157865.	PR	99US-0129845.
PR	07-OCT-1999;	99US-0158029.	PR	99US-0130077.
PR	08-OCT-1999;	99US-0158232.	PR	99US-0130449.
PR	12-OCT-1999;	99US-0158369.	PR	99US-0130510.
PR	13-OCT-1999;	99US-0159293.	PR	99US-0130891.
PR	13-OCT-1999;	99US-0159294.	PR	99US-0131449.
PR	13-OCT-1999;	99US-0159295.	PR	99US-0132048.
PR	14-OCT-1999;	99US-0159329.	PR	99US-0132407.
PR	14-OCT-1999;	99US-0159330.	PR	99US-0132484.
PR	14-OCT-1999;	99US-0159331.	PR	99US-0132485.
PR	14-OCT-1999;	99US-0159637.	PR	99US-0132486.
PR	14-OCT-1999;	99US-0159638.	PR	99US-0132487.
PR	18-OCT-1999;	99US-0159584.	PR	99US-0132863.
PR	21-OCT-1999;	99US-0160741.	PR	99US-0134256.
PR	21-OCT-1999;	99US-0160767.	PR	99US-0134218.
PR	21-OCT-1999;	99US-0160768.	PR	99US-0134219.
PR	21-OCT-1999;	99US-0160770.	PR	99US-0134221.
PR	21-OCT-1999;	99US-0160814.	PR	99US-0134370.
PR	21-OCT-1999;	99US-0160815.	PR	99US-0134768.
PR	22-OCT-1999;	99US-0160980.	PR	99US-0134941.
PR	22-OCT-1999;	99US-0160981.	PR	99US-0135124.
PR	22-OCT-1999;	99US-0160989.	PR	99US-0135353.
PR	25-OCT-1999;	99US-0161404.	PR	99US-0135629.
PR	25-OCT-1999;	99US-0161405.	PR	99US-0136021.
PR	25-OCT-1999;	99US-0161406.	PR	99US-0136392.
PR	26-OCT-1999;	99US-0161359.	PR	99US-0136782.
PR	26-OCT-1999;	99US-0161360.	PR	99US-0137222.
PR	26-OCT-1999;	99US-0161361.	PR	99US-0137528.
PR	28-OCT-1999;	99US-0161920.	PR	99US-0137502.
PR	28-OCT-1999;	99US-0161992.	PR	99US-0137724.
PR	28-OCT-1999;	99US-0161993.	PR	99US-0138094.
PR	29-OCT-1999;	99US-0162142.	PR	99US-0138540.
PR			PR	99US-0138847.
PR			PR	99US-0139119.
PR			PR	99US-0139452.
PR			PR	99US-0139453.
PR			PR	99US-0139492.
PR			PR	99US-0139454.
PR			PR	99US-0139455.
PR			PR	99US-0139456.
PR			PR	99US-0139457.
PR			PR	99US-0139458.
PR			PR	99US-0139459.
PR			PR	99US-0139460.
PR			PR	99US-0139461.
PR			PR	99US-0139462.
PR			PR	99US-0139463.
PR			PR	99US-0139750.
PR			PR	99US-0139763.
PR			PR	99US-0139817.
PR			PR	99US-0139899.
PR			PR	99US-0140353.
PR			PR	99US-0140354.
PR			PR	99US-0140695.
PR			PR	99US-0140823.
PR			PR	99US-0140991.
PR			PR	99US-0140991.

Query Match 47.08; Score 47; DB 21; Length 923;
 Best Local Similarity 66.7%; Pred. No. 99;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 KICEQIEVYFGD 17
 |||:|||||
 Db 126 KIVNQVEYFSD 137

RESULT 11
 AAG47712
 ID AAG47712 standard; Protein; 993 AA.
 AC AAG47712;
 XX
 XX 18-OCT-2000 (first entry)
 DT
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 60166.
 DE
 XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146389.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 47.0%; Score 47; DB 21; Length 993;
 Best Local Similarity 66.7%; Pred. No. 1.1e+02;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 KICEQIEYFGD 17
 || |:|||||
 Db 196 KIVNQVEYFSD 207

RESULT 12
 AAU96342
 ID AAU96342 standard; Protein; 412 AA.

XX AC AAU96342;
 XX DT 15-JUL-2002 (first entry)
 XX DE Der HMW-map polypeptide #29.
 XX Der HMW-map; American house dust mite; anti-allergic; mite; IgE;
 KW mite allergenic protein; Immunoglobulin E; hypersensitivity;
 KW immunocomplex formation.
 XX OS Dermatophagoides farinae.
 XX PN WO200222807-A2.
 XX PD 21-MAR-2002.
 XX PF 14-SEP-2001; 2001WO-US28730.

XX 14-SEP-2000; 2000US-0662293.
 XX (HESK-) HESKA CORP.
 XX McCall CA, Hunter SW, Weber ER;
 XX WPI; 2002-351888/38.
 XX N-PSDB; ABK69597.
 XX New mite allergenic protein isolated from Dermatophagoides, designated
 PT Der HMW-map protein, useful as a vaccine for treating mite allergy
 XX
 XX Example 19; Page 159-160; 161pp; English.
 XX The invention relates to an isolated mite allergenic protein of
 CC Dermatophagoides, designated Der HMW-map protein, and its related nucleic
 CC acid. The Der HMW-map protein is useful for eliciting an immune response
 CC against Der HMW-map protein. The protein or a reagent comprising a
 CC non-proteinaceous epitope is useful for identifying an animal (e.g., dog,
 CC cat) susceptible to or having an allergic response to a mite. A
 CC therapeutic composition is useful for desensitising a host animal to an
 CC allergic response to a mite. The DNA and protein can be used in the
 CC detection of anti-Der HMW-map antibodies in animal fluids, and inhibition
 CC of immunoglobulin (Ig)E or Der HMW-map protein activity associated with a
 CC disease. Antibodies that bind to Der HMW-map are useful for inhibiting
 CC binding of proteins to IgE, to prevent immunocomplex formation, thus
 CC reducing hypersensitivity responses to mite allergens, and as vaccines
 CC against mite allergen hypersensitivity. Sequences AAU96341-AAU96342
 CC represent Der HMW-map polypeptides of the invention.
 XX
 XX Sequence 412 AA;
 SQ
 Query Match 46.5%; Score 46.5; DB 23; Length 412;
 Best Local Similarity 56.2%; Pred. No. 49;
 Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
 QY 2 DLDTKICEQIEY-YFG 16
 DB :|||:| | | | |
 27 DDTSLCTHIVSYFG 42
 RESULT 13
 AAU96341
 ID AAU96341 standard; Protein; 462 AA.
 XX AAU96341;
 XX 15-JUL-2002 (first entry)
 DT Der HMW-map polypeptide #28.
 DE
 XX Der HMW-map; American house dust mite; anti-allergic; mite; IgE;
 KW mite allergenic protein; immunoglobulin E; hypersensitivity;
 KW immunocomplex formation.
 XX Dermatophagoides farinae.
 OS
 XX WO200222807-A2.
 PN
 XX 21-MAR-2002.
 PD
 XX 14-SEP-2001; 2001WO-US28730.
 PF
 XX 14-SEP-2000; 2000US-0662293.
 PR
 XX (HESK-) HESKA CORP.
 PA
 XX McCall CA, Hunter SW, Weber ER;
 PI
 XX WPI; 2002-351888/38.
 DR N-PSDB; ABK69593.
 XX

PT New mite allergenic protein isolated from Dermatophagoides, designated
 XX Der HMW-map protein, useful as a vaccine for treating mite allergy
 XX
 XX Example 19; Page 152-154; 161pp; English.
 XX The invention relates to an isolated mite allergenic protein of
 CC Dermatophagoides, designated Der HMW-map protein, and its related nucleic
 CC acid. The Der HMW-map protein is useful for eliciting an immune response
 CC against Der HMW-map protein. The protein or a reagent comprising a
 CC non-proteinaceous epitope is useful for identifying an animal (e.g., dog,
 CC cat) susceptible to or having an allergic response to a mite. A
 CC therapeutic composition is useful for desensitising a host animal to an
 CC allergic response to a mite. The DNA and protein can be used in the
 CC detection of anti-Der HMW-map antibodies in animal fluids, and inhibition
 CC of immunoglobulin (Ig)E or Der HMW-map protein activity associated with a
 CC disease. Antibodies that bind to Der HMW-map are useful for inhibiting
 CC binding of proteins to IgE, to prevent immunocomplex formation, thus
 CC reducing hypersensitivity responses to mite allergens, and as vaccines
 CC against mite allergen hypersensitivity. Sequences AAU96314-AAU96342
 CC represent Der HMW-map polypeptides of the invention.
 XX
 XX Sequence 462 AA;
 SQ
 Query Match 46.5%; Score 46.5; DB 23; Length 462;
 Best Local Similarity 56.2%; Pred. No. 55;
 Matches 9; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
 QY 2 DLDTKICEQIEY-YFG 16
 DB :|||:| | | | |
 52 DDTSLCTHIVSYFG 67
 RESULT 14
 AAU96314
 ID AAU96314 standard; Protein; 492 AA.
 XX AAU96314;
 XX 08-FEB-2001 (first entry)
 DT Human ORFX ORF2874 polypeptide sequence SEQ ID NO:5748.
 DE
 XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antihypertensive;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW neurodegenerative disorder; cancer; proliferative disorder; hypertension;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX Homo sapiens.
 OS
 XX WO200058473-A2.
 PN
 XX 05-OCT-2000.
 PD
 XX 31-MAR-2000; 2000WO-US08621.
 PF
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shimkets RA, Leach M;
 PI

XX WPI: 2000-602362/57.
 DR N-PSDB; AAC77319.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 4908-4909; 5507pp; English.
 XX
 CC AAC74446 to AAC77506 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antiproliferative; antiparkinsonian; neurotropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 492 AA;
 Query Match 46.0%; Score 46; DB 21; Length 492;
 Best Local Similarity 50.0%; Pred. No. 71;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 2 DLDPKICEQIEYFGD 17
 DB 92 ELIKLVQIEYFSD 107
 RESULT 15
 AAG17695
 ID AAG17695 standard; Protein: 223 AA.
 AC AAG17695;
 DT 17-OCT-2000 (first entry)
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 18813.
 DE Arabidopsis thaliana.
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 XX EP1033405-A2.
 PN 06-SEP-2000.
 PD 25-FEB-2000; 2000EP-0301439.
 PF 25-FEB-1999; 99US-0121825.
 XX 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 23-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 03-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140931.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.

PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144332.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145911.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	10-AUG-1999;	99US-0147935.
PR	11-AUG-1999;	99US-0148171.
PR	12-AUG-1999;	99US-0148319.
PR	13-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	16-AUG-1999;	99US-0148684.
PR	17-AUG-1999;	99US-0149368.
PR	18-AUG-1999;	99US-0149175.
PR	20-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157765.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:34:48 ; Search time 20.0225 Seconds
(without alignments)
185.234 Million cell updates/sec

Title: US-09-836-073-15

Perfect score: 100

Sequence: 1 LDLTKICEIYFYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 9297

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	33	33.0	24	2	Q05616		Q05616 staphylococ
2	29	29.0	16	2	Q9R963		Q9R963 helicobacte
3	28	28.0	11	4	Q9UC46		Q9UC46 homo sapien
4	28	28.0	23	4	Q9UC18		Q9UC18 homo sapien
5	28	28.0	25	8	Q9XJ27		Q9XJ27 aureomembra
6	27.5	27.5	24	5	Q9BM09		Q9BM09 spongilla 1
7	27	27.0	18	4	Q16028		Q16028 homo sapien
8	27	27.0	20	4	Q9UC68		Q9UC68 homo sapien
9	26	26.0	16	2	P83153		P83153 anaabaena sp
10	25.5	25.5	24	9	Q38270		Q38270 bacterioph
11	25	25.0	19	2	Q9R4B9		Q9R4B9 streptococc
12	25	25.0	20	6	P79256		P79256 actus trivi
13	25	25.0	20	8	Q9TGB8		Q9TGB8 alnus crisp
14	25	25.0	25	8	Q9TGB7		Q9TGB7 alnus glut
15	25	25.0	25	8	Q9TGB6		Q9TGB6 alnus marit
16	25	25.0	25	8	Q9TGB5		Q9TGB5 betula alle

17	25	25.0	25	8	Q9TGB4	Q9TGB4 betula glan
18	25	25.0	25	8	Q9TGB3	Q9TGB3 betula papy
19	25	25.0	25	8	Q9TGB2	Q9TGB2 betula verr
20	25	25.0	25	8	Q9TGB1	Q9TGB1 betula pube
21	25	25.0	25	8	Q9TGB0	Q9TGB0 corylus ave
22	25	25.0	25	8	Q9TGA9	Q9TGA9 corylus col
23	25	25.0	25	8	Q9TGA8	Q9TGA8 corylus cor
24	25	25.0	25	8	Q9TGA7	Q9TGA7 ostrya virg
25	25	25.0	25	8	Q9TGA6	Q9TGA6 quercus rub
26	25	25.0	25	8	Q9TGB3	Q9TGB3 carpinus ca
27	24.5	24.5	20	5	P82109	P82109 mythinna un
28	24	24.0	13	8	Q9THR8	Q9THR8 bryopsis sp
29	24	24.0	16	2	Q45663	Q45663 bacillus su
30	24	24.0	19	2	Q47895	Q47895 fremyella d
31	24	24.0	20	10	Q9S876	Q9S876 triticum ae
32	24	24.0	23	4	Q9S8B3	Q9S8B3 homo sapien
33	23.5	23.5	14	11	Q921H4	Q921H4 mus musculu
34	23	23.0	12	10	Q9M433	Q9M433 lotus japon
35	23	23.0	15	9	Q9MBH6	Q9MBH6 streptococc
36	23	23.0	17	13	Q91352	Q91352 gallus gall
37	23	23.0	19	10	P83089	P83089 spinacia ol
38	23	23.0	20	5	Q9TWR0	Q9TWR0 biattella g
39	23	23.0	20	10	Q9S8X5	Q9S8X5 glycine max
40	23	23.0	23	11	Q9QV93	Q9QV93 rattus sp.
41	23	23.0	25	5	Q9BM63	Q9BM63 lissomyema
42	23	23.0	25	13	Q9PRR5	Q9PRR5 anser anser
43	23	23.0	25	15	Q70139	Q70139 human immun
44	22.5	22.5	21	2	Q93C14	Q93C14 escherichia
45	22	22.0	12	2	Q56947	Q56947 versinia ps

ALIGNMENTS

RESULT 1
Q05616 PRELIMINARY; PRT; 24 AA.
ID AC Q05616;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE 3-dehydroquinase synthase (EC 4.6.1.3) (Fragment).
GN AROB.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8325-4;
RX MEDLINE=93381456; PubMed=8371108;
RA O'Connell C.M., Pattee P., Foster T.J.;
RT "Sequence and mapping of the *aroA* gene of *Staphylococcus aureus* 8325-4";
RL J. Gen. Microbiol. 139:1449-1460(1993).
CC -!- CATALYTIC ACTIVITY: 7-PHOSPHO-3-DEOXY-ARABINO-HEPTULOSONATE - 3-DEHYDROQUINATE + ORTHOPHOSPHATE.
CC -!- PATHWAY: SECOND STEP IN THE BIOSYNTHESIS FROM CHORISMATE OF THE AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
DR EMBL; L05004; AAA71896.1; -;
KW Aromatic amino acid biosynthesis; Lyase.
FT NON_TER 1
SQ SEQUENCE 24 AA; 2857 MW; 1C4C8D20F7A4856E CRC64;
Query Match 33.0%; Score 33; DB 2; Length 24;
Best Local Similarity 62.5%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 8 CEQIEYF 15
DB 16 CEQLKTYF 23

```

RESULT 2
Q9R963
ID Q9R963 PRELIMINARY; PRT; 16 AA.
AC Q9R963;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE FepC (Fragment).
GN FepC
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
ON NCBI_TaxID=210;
RX STRAIN=F31;
RP SEQUENCE FROM N.A.
RC MEDLINE=98453456; PubMed=9780260;
RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,
RA Kuriyama M.;
RT "Full-length sequence analysis of the vacA gene from cytotoxic and
RT noncytotoxic Helicobacter pylori."
RL J. Infect. Dis. 178:1391-1398(1998).
DR ENBL: AF049623; AAD04263.1; -
FT NON_TER 1
FT NON_TER 16
SQ SEQUENCE 16 AA; 1938 MW; 3C33D403ABE61428 CRC64;

Query Match 29.0%; Score 29; DB 2; Length 16;
Best Local Similarity 55.6%; Pred. No. 6.9e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYFGDF 18
   : : : : :
Db 7 QVEYAFNF 15

RESULT 3
Q9UC46
ID Q9UC46 PRELIMINARY; PRT; 11 AA.
AC Q9UC46;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Neutrophil inhibitor peptide, NIP-POLYMORPHONUCLEAR neutrophil
DE inhibitor peptide.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=96326114; PubMed=8703476;
RA Cooper J.A.Jr., Culbreth R.R.;
RT "Characterization of a neutrophil inhibitor peptide harvested from
RT human bronchial lavage: homology to influenza A nucleoprotein."
RL Am. J. Respir. Cell Mol. Biol. 15:207-215(1996).
SQ SEQUENCE 11 AA; 1262 MW; 951A1C3279C9DB45 CRC64;

Query Match 28.0%; Score 28; DB 4; Length 11;
Best Local Similarity 80.0%; Pred. No. 6.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 13 YFGD 17
   : : : :
Db 5 YFGD 9

RESULT 4
Q9UCL8
ID Q9UCL8 PRELIMINARY; PRT; 23 AA.
AC Q9UCL8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE GLYCOSAMINOGLYCAN-BEARING polypeptide (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=93098810; PubMed=1463459;
RA Bonnet F., Perin J.P., Mailliet P., Jolles P., Alliel P.M.;
RT "Characterization of a human seminal plasma glycosaminoglycan-bearing
RT polypeptide."
RL Biochem. J. 288:565-569(1992).
SQ SEQUENCE 23 AA; 2334 MW; E469397E6B2EA794 CRC64;

Query Match 28.0%; Score 28; DB 4; Length 23;
Best Local Similarity 45.5%; Pred. No. 1.4e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 CEQIEYFGDF 18
   : : : : :
Db 7 CEQEQTGXGDF 17

RESULT 5
Q9XJZ7
ID Q9XJZ7 PRELIMINARY; PRT; 25 AA.
AC Q9XJZ7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit
DE (Fragment).
GN RBCS.
OS Aureocymbra lagunensis.
OG Chloroplast.
OC Eukaryota; stramenopiles; Pelagophyceae; Aureocymbra.
OX NCBI_TaxID=44058;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCMP1681, CCMP1509, AND CCMP1510;
RA Bailey J.C., Andersen R.A.;
RT "Analysis of clonal cultures of the brown tide algae Aureococcus and
RT Aureocymbra (Pelagophyceae) using 18S rRNA, rbcL and rubisco spacer
RT sequences."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR ENBL: AF117782; AAD39106.1; -
DR ENBL: AF117780; AAD39104.1; -
DR ENBL: AF117781; AAD39105.1; -
DR Interpro: IPR000894; RuBisCO_small.
DR Pfam: PF00101; RuBisCO_small; 1.
KW Chloroplast.
FT NON_TER 25
SQ SEQUENCE 25 AA; 2972 MW; 61A96BD83DB94118 CRC64;

Query Match 28.0%; Score 28; DB 8; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 DTKICEQIEY 13
   : : : : :
Db 16 DDQIRKQVEY 25

RESULT 6
Q9BM09
ID Q9BM09 PRELIMINARY; PRT; 24 AA.
AC Q9BM09;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Gypsy-like reverse transcriptase (Fragment).

```

```
OS Spongilla lacustris (Freshwater sponge);
OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;
OC Haplosclerida; Spongillidae; Spongilla.
OX NCBI_TaxID=6055;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-GRT-G7 RETROTRANSPOSON;
RX MEDLINE=20570504; PubMed=11121049;
RA Arkhipova I., Meselson M.;
RT "Transposable elements in sexual and ancient asexual taxa.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
DR EMBL; AY013997; AAG59969.1;
KW RNA-directed DNA polymerase.
FT NON_TER 1
FT NON_TER 24
SQ SEQUENCE 24 AA; 2754 MW; B588512EDE7A251A CRC64;

Query Match 27.5%; Score 27.5; DB 5; Length 24;
Best Local Similarity 38.5%; Pred. No. 1.8e+03;
Matches 5; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 3 LTKICEQIEYF 15
Db 12 MDT-VCHGLEFVF 23

RESULT 7
Q16028
ID Q16028 PRELIMINARY; PRT; 18 AA.
AC Q16028;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE OCRL-1 protein (Fragment).
GN OCRL-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93278398; PubMed=8504307;
RA Leachey A.M., Charnas L.R., Nusbaum R.L.;
RT "Nonense mutations in the OCRL-1 gene in patients with the
RT Oculocerebrorenal syndrome of Lowe.";
RL Hum. Mol. Genet. 2:461-463(1993).
DR EMBL; S62084; AAD13933.1;
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 18 AA; 2055 MW; 61356C825F7AF33D CRC64;

Query Match 27.0%; Score 27; DB 4; Length 18;
Best Local Similarity 50.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 DTKICEQI 11
Db 4 DPRICRQL 11

RESULT 8
Q9UCB8
ID Q9UCB8 PRELIMINARY; PRT; 20 AA.
AC Q9UCB8;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE D(TAGG)N-binding protein B39 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
```

```
RX MEDLINE=93309464; PubMed=8321232;
RA Ishikawa F., Matunis M.J., Dreyfuss G., Cech T.R.;
RT "Nuclear proteins that bind the pre-mRNA 3' splice site sequence
RT r(UUAG/G) and the human telomeric DNA sequence d(TTAGGG)n.";
RL Mol. Cell. Biol. 13:4301-4310(1993).
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 1.
SQ SEQUENCE 20 AA; 2255 MW; C3CE1955E9A6D210 CRC64;

Query Match 27.0%; Score 27; DB 4; Length 20;
Best Local Similarity 62.5%; Pred. No. 1.8e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 9 EOIEYFYG 16
Db 13 EKIREYFG 20

RESULT 9
P83153
ID P83153 PRELIMINARY; PRT; 16 AA.
AC P83153;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE C-phycocyanin beta chain (Fragment).
OS Anabaena sp. (strain L31).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=29412;
RN [1]
RP SEQUENCE.
RA Apte S.K., Uhlemann E., Schmid R., Altendorf K.;
RL Submitted (OCT-2001) to the SWISS-PROT data bank.
CC -!- FUNCTION: LIGHT-HARVESTING PHOTOSYNTHETIC BILE PIGMENT-PROTEIN
CC FROM THE PHYCOBILIPROTEIN COMPLEX.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
CC -!- SUBCELLULAR LOCATION: THYLAKOID MEMBRANE.
CC -!- PTM: CONTAINS TWO COVALENTLY LINKED BILIN CHROMOPHORES (BY
CC SIMILARITY).
KW Phycobilisome; Electron transport; Photosynthesis; Bile pigment;
KW Thylakoid; Membrane.
FT NON_TER 16
SQ SEQUENCE 16 AA; 1723 MW; C272F2E8311CCF9E CRC64;

Query Match 26.0%; Score 26; DB 2; Length 16;
Best Local Similarity 41.7%; Pred. No. 2.1e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLDTRKICEQIE 12
Db 2 LDVTRKVSQAD 13

RESULT 10
Q38270
ID Q38270 PRELIMINARY; PRT; 24 AA.
AC Q38270;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Bacteriophage lambda mutant truncated regulatory protein (PyrI) gene
DE from E.coli, 3' end (Fragment).
OS Bacteriophage lambda.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=10710;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86143826; PubMed=3912513;
RA Cunin R., Jacobs A., Charlier D., Crabeel M., Herve G., Glansdorff N.,
RA Pierard A.;
RT "Structure-function relationship in allosteric aspartate
RT carbamoyltransferase from Escherichia coli: I. Primary structure of a
```


RT pyrI gene encoding a modified regulatory subunit.";

RL J. Mol. Biol. 186:707-713(1985).

DR EMBL; M28579; AAA32252.1; "

DR HSSP; P00478; 8ATC.

FT NON_TER 1

SQ SEQUENCE 24 AA; 2905 MW; FD9349DF6F9159D0 CRC64;

Query Match 25.5%; Score 25.5; DB 9; Length 24;

Best Local Similarity 50.0%; Pred. No. 3.7e+03;

Matches 6; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 3 LDTKICEQIEY 14

Db 9 LKCKYCEK-EFY 19

RESULT 11

Q9R4B9 ID Q9R4B9 PRELIMINARY; PRT; 19 AA.

AC Q9R4B9;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

DE Surface trypsin-resistant LADDERING protein (Fragment).

OS Streptococcus sp.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;

OC Streptococcaceae; Streptococcus.

OX NCBI_TaxID=1306;

RN [1]

RP SEQUENCE.

RX MEDLINE=97047708; PubMed=8926097;

RA Lachenauer C.S., Madoff L.C.;

RT "A protective surface protein from type V group B streptococci shares

RT N-terminal sequence homology with the alpha C protein.";

RL Infect. Immun. 64:4255-4260(1996).

SQ SEQUENCE 19 AA; 2008 MW; F44CC89ED774EA3C CRC64;

Query Match 25.0%; Score 25; DB 2; Length 19;

Best Local Similarity 50.0%; Pred. No. 3.5e+03;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LDTKICEQIE 12

Db 9 LNTKITKNIQ 18

RESULT 12

F79256 ID F79256 PRELIMINARY; PRT; 20 AA.

AC P79256;

DT 01-MAY-1997 (TrEMBLrel. 03, Created)

DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)

DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)

DE S-cone pigment gene/opsin protein (Fragment).

OS Aotus trivirgatus (Night monkey) (Douroucouli).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.

OX NCBI_TaxID=9505;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96288974; PubMed=8763792;

RA Jacobs G.H., Neitz M., Neitz J.;

RT "Mutations in S-cone pigment genes and the absence of colour vision in

RT two species of nocturnal primate."

RL Proc. R. Soc. Lond., B, Biol. Sci. 263:705-710(1996).

DR EMBL; S82618; AAD14408.1; "

FT NON_TER 20

SQ SEQUENCE 20 AA; 2325 MW; 1F017A36C0C4101E CRC64;

Query Match 25.0%; Score 25; DB 6; Length 20;

Best Local Similarity 40.0%; Pred. No. 3.7e+03;

Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 5 TKICEQIEY 14

Db 2 SKMSEEEFY 11

RESULT 13

Q9TGB8 ID Q9TGB8 PRELIMINARY; PRT; 25 AA.

AC Q9TGB8;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Ribosomal protein small 3 (Fragment).

GN RPS3.

OS Alnus crispa.

OG Mitochondrion.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids I; Fagales; Betulaceae; Alnus.

OX NCBI_TaxID=3518;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99261656; PubMed=10331271;

RA Laroche J., Bousquet J.;

RT "Evolution of the mitochondrial rps3 intron in perennial and annual

RT angiosperms and homology to nad5 intron 1.";

RL Mol. Biol. Evol. 16:441-452(1999).

DR EMBL; AF080075; AAD50062.1; "

DR InterPro; IPR001351; Ribosomal_S3.

DR Pfam; PF00417; Ribosomal_S3_N; 1.

KW Mitochondrion.

FT NON_TER 1

FT NON_TER 25

SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 25.0%; Score 25; DB 8; Length 25;

Best Local Similarity 60.0%; Pred. No. 4.6e+03;

Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 12 EYFYG 16

Db 4 DYYYG 8

RESULT 14

Q9TGB7 ID Q9TGB7 PRELIMINARY; PRT; 25 AA.

AC Q9TGB7;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Ribosomal protein small 3 (Fragment).

GN RPS3.

OS Alnus glutinosa (Alder).

OG Mitochondrion.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids I; Fagales; Betulaceae; Alnus.

OX NCBI_TaxID=3517;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99261656; PubMed=10331271;

RA Laroche J., Bousquet J.;

RT "Evolution of the mitochondrial rps3 intron in perennial and annual

RT angiosperms and homology to nad5 intron 1.";

RL Mol. Biol. Evol. 16:441-452(1999).

DR EMBL; AF080076; AAD50063.1; "

DR InterPro; IPR001351; Ribosomal_S3.

DR Pfam; PF00417; Ribosomal_S3_N; 1.

KW Mitochondrion.

FT NON_TER 1

FT NON_TER 25

SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 25.0%; Score 25; DB 8; Length 25;
 Best Local Similarity 60.0%; Pred. No. 4.6e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
 Db 4 DYYG 8

RESULT 15

Q9TGB6 PRELIMINARY; PRT; 25 AA.
 AC Q9TGB6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Ribosomal protein small 3 (Fragment).
 GN RPS3.
 OS Alnus maritima.
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fagales; Betulaceae; Alnus.
 OX NCBI_TaxID=21015;
 RN [1]
 RP SEQUENCE FROM N.A. PubMed=10331271;
 RX MEDLINE=99261656; PubMed=10331271;
 RA Laroche J., Bousquet J.;
 RT "Evolution of the mitochondrial rps3 intron in perennial and annual
 angiosperms and homology to nad5 intron 1".
 RL Mol. Biol. Evol. 16:441-452(1999).
 DR EMBL: AF080077; AAD50064.1; -;
 DR InterPro: IPR001351; Ribosomal_S3.
 DR Pfam: PF00417; Ribosomal_S3_N; 1.
 KW Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 25 25
 SQ SEQUENCE 25 AA; 3084 MW; 3CAA2B3D962A52D8 CRC64;

Query Match 25.0%; Score 25; DB 8; Length 25;
 Best Local Similarity 60.0%; Pred. No. 4.6e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFG 16
 Db 4 DYYG 8

Search completed: April 23, 2003, 13:47:16
 Job time : 21.0225 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:33:13 ; Search time 4.65169 Seconds
(without alignments)
160.495 Million cell updates/sec

Title: US-09-836-073-15

Perfect score: 100

Sequence: 1 LLDLTKICEQIEYVFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1520

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	31.0	23	1	XYC1_ACIGB
2	30	30.0	18	1	PHT_PSESE
3	27	27.0	19	1	HBB2_UROHA
4	27	27.0	23	1	PRO3_DACGL
5	25	25.0	18	1	MLB_SCYCA
6	24	24.0	16	1	MLB_SQUAC
7	23	23.0	24	1	VPB_BPDI0
8	22	22.0	13	1	AU11_LITRA
9	22	22.0	23	1	AFP2_BRANA
10	22	22.0	25	1	ANDT_ANDAU
11	21	21.0	14	1	LECB_PSOSC
12	21	21.0	23	1	NUO5_SOLTU
13	21	21.0	23	1	UDP_LACCA
14	21	21.0	24	1	PQQA_ACICA
15	21	21.0	24	1	PQQA_PSEFL
16	21	21.0	25	1	SMBP_RAT
17	20	20.0	18	1	SFAH_HELAN
18	20	20.0	20	1	COGA_CHIOP
19	20	20.0	20	1	FIBB_FELCA
20	20	20.0	20	1	PSAL_SYNNU
21	20	20.0	20	1	SPTG_PSEUS
22	19	19.5	24	1	GAE6_RANRU
23	19	19.0	8	1	ALI7_CARMA
24	19	19.0	9	1	ALI1_LEUMA
25	19	19.0	11	1	LSK1_SOYBN
26	19	19.0	13	1	ACT7_SOYBN
27	19	19.0	15	1	LPE_ECOLI
28	19	19.0	16	1	DHE2_THUTH
29	19	19.0	17	1	ACT6_SOYBN
30	19	19.0	17	1	GAST_MACMU
31	19	19.0	19	1	FIBB_VULVU
32	19	19.0	20	1	COXN_THUOB
33	19	19.0	20	1	OAR_PHOPY

34	19	19.0	20	1	PGK_CLOPA
35	19	19.0	20	1	THIO_CANFA
36	19	19.0	20	1	TL22_SPIOL
37	19	19.0	23	1	COXK_CANFA
38	19	19.0	23	1	CR42_LITCE
39	19	19.0	24	1	LPTN_ECOLI
40	18	18.0	4	1	OCPL_OCTMI
41	18	18.0	8	1	LMT2_LOCHI
42	18	18.0	9	1	FLA2_TREHY
43	18	18.0	12	1	NO40_SESRO
44	18	18.0	13	1	AUI2_LITRA
45	18	18.0	13	1	NPI_LYMST

ALIGNMENTS

RESULT 1
XYC1_ACIGB
ID XYC1_ACIGB STANDARD; PRT; 23 AA.
AC P46365;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Benzaldehyde dehydrogenase [NAD+] I (EC 1.2.1.28) (Fragment).
OS Acinetobacter genosp. 11.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=106649;
RN [1]
RP SEQUENCE.
RC STRAIN=ATCC 11171 / NCIB 8250 / CIP 63.46 / B94;
RA MEDLINE=91113163; PubMed=1989592;
RA Chalmers R.M., Keen J.N., Fawson C.A.;
RT "Comparison of benzyl alcohol dehydrogenases and benzaldehyde dehydrogenases from the benzyl alcohol and mandelate pathways in Acinetobacter calcoaceticus and from the TOL-plasmid-encoded toluene pathway in Pseudomonas putida. N-terminal amino acid sequences, amino acid compositions and immunological cross-reactions.";
RL Biochem. J. 273:199-107(1991).
CC -|- CATALYTIC ACTIVITY: Benzaldehyde + NAD(+) + H(2)O = benzoate + NADH.
CC -|- SUBUNIT: HOMOTETRAMER.
CC -|- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
CC InterPro: IPR002086; Aldehyde dehydr.
DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS; PARTIAL.
DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU; PARTIAL.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2750 MW; 5049604930DF9E40 CRC64;

Query Match 31.0%; Score 31; DB 1; Length 23;
Best Local Similarity 50.0%; Pred. No. 60;
Matches 8; Conservative 3; Mismatches 3; Indels 2; Gaps 1;
QY 2 DLDTKICEQI--EYVF 15
:: ||||| |:: |
Db 2 NIQTKEIQWKEHIF 17

RESULT 2
PHT_PSESE
ID PHT_PSESE STANDARD; PRT; 18 AA.
AC P25271;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-MAY-1992 (Rel. 22, Last annotation update)
DE Phormonotropin (PSS-PT).
OS Pseudaletia separata (Armyworm).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC Ditrysia; Noctuoidea; Noctuidae; Hadeninae; Pseudaletia.
OX NCBI_TaxID=7105;

```

RN  SEQUENCE.
RP  TISSUE-Head;
RC  MEDLINE=92134266; PubMed=1734867;
RA  Matsumoto S., Fonagy A., Kurihara M., Uchiumi K., Nagamine T.,
RA  Chijimatsu M., Mitsui T.;
RT  "Isolation and primary structure of a novel pheromonotropic
RT  neuropeptide structurally related to leucopyrokinin from the armyworm
RT  larvae, Pseudaletia separata.";
RL  Biochem. Biophys. Res. Commun. 182:534-539(1992).
CC  -|- FUNCTION: HORMONE THAT CONTROLS SEX PHEROMONE PRODUCTION IN
CC  LARVAE.
CC  -|- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR  PIR: JS0647; JS0647.
DR  InterPro: IPR001484; Pyrokinin.
DR  PROSITE: PS00339; PYROKININ; 1.
KW  Hormone; Neuropeptide; Amidation; Pyrokinin.
FT  MOD_RES 18 18
SQ  SEQUENCE 18 AA; 2200 MW; DDF3A1B75ACB18FF CRC64;

Query Match 30.0%; Score 30; DB 1; Length 18;
Best Local Similarity 38.5%; Pred. No. 69;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LLDLTKICEQIEY 13
Db 1 LLDLTKICEQIEY 13
2 LSYDDKVFENVEF 14
2 LSYDDKVFENVEF 14

RESULT 3
HBB2_UROHA
ID HBB2_UROHA STANDARD; PRT; 19 AA.
AC P18992;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hemoglobin beta-2 chain (Fragment).
OS Uromastix hardwickii (Indian spiny-tailed lizard).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Agamidae; Uromastycinae;
OC Uromastix.
OX NCBI_TaxID=40250;
RN [1]
RP SEQUENCE.
RX MEDLINE=84029159; PubMed=6628672;
RA Naqvi S., Zaidi Z.H., von Bahr-Lindstroem H., Carlquist M.,
RA Joernvall H.;
RT "Characterization of hemoglobin from the lizard Uromastix
RT hardwickii.";
RL FEBS Lett. 162:290-295(1983).
CC -|- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE
CC VARIOUS PERIPHERAL TISSUES..
CC -|- SUBUNIT: HETEROTETRAMER OF TWO ALPHA CHAINS AND TWO BETA CHAINS.
CC -|- TISSUE SPECIFICITY: RED BLOOD CELLS.
CC -|- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR PIR: A05305; A05305.
DR InterPro: IPR00371; Globin.
DR PROSITE: PS01033; GLOBIN; PARTIAL.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT NON_TER 1 1
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 1914 MW; C40AD8EA30019057 CRC64;

Query Match 27.0%; Score 27; DB 1; Length 19;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 YFGDF 18
Db 1 YFGDF 5

RESULT 4

```

```

PRO3_DACGL
ID PRO3_DACGL STANDARD; PRT; 23 AA.
AC P18690;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Prolamin alpha-3 (Fragment).
OS Dactylis glomerata (Orchard grass) (Cocksfoot grass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poaceae; Dactylis.
OX NCBI_TaxID=4509;
RN [1]
RP SEQUENCE.
RA Vvedenskaya I.O., Shlyapnikov S.V., Konarev A.V.;
RT "Characterization of the N-terminal amino acid sequence of alpha-
RT prolamine from Dactylis glomerata L.";
RL Biochimia 51:1519-1522(1986).
DR PIR: S02201; S02201.
DR NON_TER 23
SQ SEQUENCE 23 AA; 2938 MW; E756CA1217592F39 CRC64;

Query Match 27.0%; Score 27; DB 1; Length 23;
Best Local Similarity 42.9%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 LLDLTKICEQIEY 14
Db 1 LLDLTKICEQIEY 14
2 VQDDPFQEQQY 15
2 VQDDPFQEQQY 15

RESULT 5
MLB_SCYCA
ID MLB_SCYCA STANDARD; PRT; 18 AA.
AC P01206;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanotropin beta (Beta-MSH).
OC Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC Scyliorhinidae; Scyliorhinus.
OX NCBI_TaxID=7830;
RN [1]
RP SEQUENCE.
RX MEDLINE=75113445; PubMed=4452470;
RA Love R.M., Pickering B.T.;
RT "A beta-MSH in the pituitary gland of the spotted dogfish
RT (Scyliorhinus canicula): isolation and structure.";
RL Gen. Comp. Endocrinol. 24:398-404(1974).
CC -|- SIMILARITY: BELONGS TO THE POMC FAMILY.
DR PIR: A01470; MTDPEC.
KW Hormone.
SQ SEQUENCE 18 AA; 2195 MW; 353C93A30C314DF4 CRC64;

Query Match 25.0%; Score 25; DB 1; Length 18;
Best Local Similarity 40.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 9 EQIEYFGDF 18
Db 1 BZBYKMGHF 10

RESULT 6
MLB_SQUAC
ID MLB_SQUAC STANDARD; PRT; 16 AA.
AC P01207;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanotropin beta.

```

OS Squalus acanthias (Spiny dogfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
 OX NCBI_TaxID=7797;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=75127390; PubMed=4375978;
 RA Bennett H.P.J., Lowy P.J., McMartin C., Scott A.P.;
 RT "Structural studies of alpha-melanocyte-stimulating hormone and a
 RT novel beta-melanocyte-stimulating hormone from the neurointermediate
 RT lobe of the pituitary of the dogfish Squalus acanthias.";
 RL Biochem. J. 141:439-444(1974).
 CC -1- SIMILARITY: BELONGS TO THE POMC FAMILY.
 DR PIR: A01471; MTFBFS.
 KW Hormone.
 SQ SEQUENCE 16 AA; 1939 MW; 993AF43C8A640A0E CRC64;
 Query Match 24.0%; Score 24; DB 1; Length 16;
 Best Local Similarity 57.1%; Pred. No. 5.8e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 12 EYFGDF 18
 Db :| | | |
 4 DYKFGHF 10
 RESULT 7
 VBP_BPD10
 ID VBP_BPD10 STANDARD; PRT; 24 AA.
 AC P13771;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-NOV-1991 (Rel. 20, Last annotation update)
 DE DNA transposition protein (Protein B) (Fragment).
 GN KIL.
 OS Bacteriophage D108.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC Mu-like viruses.
 OX NCBI_TaxID=10671;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88226005; PubMed=2836266;
 RA Waggoner B.T., Wade T., Pato M.L.;
 RT "Identification of the bacteriophage D108 kil gene and of the second
 RT region of sequence nonhomology with bacteriophage Mu.";
 RL Gene 62:111-119(1988).
 CC -1- FUNCTION: THIS PROTEIN IS AN ESSENTIAL GENE PRODUCT THAT IS
 CC INVOLVED IN BACTERIOPHAGE INTEGRATION AND REPLICATION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M18902; AAA32205.1;
 KW DNA-binding; DNA replication; Transposition.
 FT NON_TER 1
 SQ SEQUENCE 24 AA; 2909 MW; 136DA5DDD5C6D740 CRC64;
 Query Match 23.0%; Score 23; DB 1; Length 24;
 Best Local Similarity 71.4%; Pred. No. 1.3e+03;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 LDLTKI 7
 Db | | | | |
 12 LDLVDI 18
 RESULT 8
 AUI1_LITRA

ID AUI1_LITRA STANDARD; PRT; 13 AA.
 AC P82386;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Aurein 1.1.
 OS Litoria raniformis (Southern bell frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
 OC Pelodyadinae; Litoria.
 OX NCBI_TaxID=116057;
 RN [1]
 RP SEQUENCE, AND FUNCTION.
 RC TISSUE=Granular dorsal gland;
 RX MEDLINE=20408845; PubMed=10951191;
 RA Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
 RA Wallace J.C., Tyler M.J.;
 RT "The antibiotic and anticancer active aurein peptides from the
 RT Australian bell frogs Litoria aurea and Litoria raniformis the
 RT solution structure of aurein 1.2.";
 RL Eur. J. Biochem. 267:5330-5341(2000).
 CC -1- FUNCTION: ANTIMICROBIAL ACTIVITY AGAINST B.CEREUS, L.LACTIS,
 CC L.INNOCUA AND S.UBERIS. PROBABLY ACTS BY DISTURBING MEMBRANE
 CC FUNCTIONS WITH ITS AMPHIPATHIC STRUCTURE.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC Antibiotic; Amidation.
 KW MOD_RES 13
 FT 13
 SQ SEQUENCE 13 AA; 1447 MW; 173CB99DFBC83330 CRC64;
 Query Match 22.0%; Score 22; DB 1; Length 13;
 Best Local Similarity 50.0%; Pred. No. 1e+03;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 2 DLDTKICEQI 11
 Db :| | | | |
 4 DIKKIAESI 13
 RESULT 9
 AFP2_BRANA
 ID AFP2_BRANA STANDARD; PRT; 23 AA.
 AC P30226;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cysteine-rich antifungal protein 2 (AFP2) (Fragment).
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Seed;
 RX MEDLINE=93138130; PubMed=8422949;
 RA Tertas F.R.G., Torrekens S., van Leuven F., Osborn R.W.,
 RA Vanderleyden J., Cammue B.P.A., Broekaert W.F.;
 RT "A new family of basic cysteine-rich plant antifungal proteins from
 RT Brassicaceae species.";
 RL FEBS Lett. 316:233-240(1993).
 CC -1- FUNCTION: POSSESSES ANTIFUNGAL ACTIVITY SENSITIVE TO INORGANIC
 CC CATIONS.
 CC -1- SUBUNIT: FORMS OLIGOMERS IN ITS NATIVE STATE.
 CC -1- SIMILARITY: BELONGS TO THE PLANT DEFENSIN FAMILY.
 DR PIR: S28992; S28992.
 DR HSP; P30231; IAYJ.
 DR InterPro: IPR002118; Gamma-thionin.
 DR ProDom: PD002594; Gamma-thionin; 1.
 DR PROSITE: PS00940; GAMMA_THIONIN; PARTIAL.
 KW Plant defense; Fungicide.
 FT MOD_RES 1
 FT NON_TER 23
 SQ SEQUENCE 23 AA; 2467 MW; A11D4A9E364F5735 CRC64;

Query Match 22.0%; Score 22; DB 1; Length 23;
 Best Local Similarity 60.0%; Pred. No. 1.8e+03;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 KICQ 10
 DB 2 KLCR 6

RESULT 10
 ANDT_ANDAU STANDARD; PRT; 25 AA.
 AC P36684; P81616;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUN-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Androctonus.
 OS Androctonus australis hector (Sahara scorpion).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Buthoidea; Buthidae; Androctonus.
 OX NCBI_TaxID=70175;
 RN [1]
 RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
 RC TISSUE-Hemolymph;
 RX MEDLINE=97094646; PubMed=8939880;
 RA Ehret-Sabatier L., Loew D., Goyffon M., Fehlbaum P., Hoffmann J.A.,
 van Dorselaer A., Bulet P.;
 RT "Characterization of novel cysteine-rich antimicrobial peptides from
 scorpion blood";
 RL J. Biol. Chem. 271:29537-29544(1996).
 RN [2]
 RP SYNTHESIS OF D-AMINO ACID ENANTIOMER, AND CHARACTERIZATION.
 RX MEDLINE=20115101; PubMed=10642525;
 RA Hetru C., Letellier L., Oren Z., Hoffmann J.A., Shai Y.;
 RT "Androctonin, a hydrophilic disulphide-bridged non-haemolytic
 anti-microbial peptide: a plausible mode of action";
 RL Biochem. J. 345:653-664(2000).
 RN [3]
 RP STRUCTURE BY NMR.
 RX MEDLINE=20025109; PubMed=10563585;
 RA Mandard N., Sy D., Maufrais C., Bonmatin J.M., Bulet P., Hetru C.,
 Vovelle F.;
 RT "Androctonin, a novel antimicrobial peptide from scorpion Androctonus
 australis: solution structure and molecular dynamics simulations in
 the presence of a lipid monolayer";
 RL J. Biomol. Struct. Dyn. 17:367-380(1999).
 CC -1- FUNCTION: ACTIVE AGAINST BOTH BACTERIA (GRAM-POSITIVE AND GRAM-
 NEGATIVE) AND FILAMENTOUS FUNGI. ACTS ON THE MEMBRANE OF THE
 BACTERIAL CELLS. IT DESTABILIZE A MEMBRANE BY MODIFYING ITS
 PROPERTIES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MASS SPECTROMETRY: MW=3076.7; METHOD=Electrospray.
 DR PDB: 1CZ6; 12-JAN-00.
 KW Antibiotic; Fungicide; 3D-structure.
 FT DISULFID 4 20
 FT DISULFID 10 16
 SQ SEQUENCE 25 AA; 3081 MW; D4183D6EA5A50AA CRC64;

Query Match 22.0%; Score 22; DB 1; Length 25;
 Best Local Similarity 50.0%; Pred. No. 2e+03;
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICEQIE 12
 DB 3 VCRQIK 8

RESULT 11
 LECB_PSOSC STANDARD; PRT; 14 AA.
 AC P22584;
 DT 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Basic lectin B1 (Fragment).
 OS Psophocarpus scandens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 OC Psophocarpus.
 OX NCBI_TaxID=3890;
 RN [1]
 RN SEQUENCE.
 RP TISSUE=Seed;
 RC Kott A.A.;
 RT "Isolation and characterization of the lectins from the seeds of
 Psophocarpus scandens";
 RL Phytochemistry 27:2847-2855(1988).
 CC -1- SUBUNIT: DIMER OF TWO IDENTICAL OR NEARLY IDENTICAL SUBUNITS OF
 ABOUT 32000 APPARENT MW.
 CC -1- PTM: CONTAINS ABOUT 9% CARBOHYDRATE. HETEROGENEITY IN CARBOHYDRATE
 COMPOSITION MAY RESULT IN DIFFERENT ISOLECTIN FORMS.
 CC -1- SIMILARITY: WITH P.TETRAGONOLBUS BASIC LECTINS IN N-TERMINAL
 SEQUENCE, AA COMPOSITION, ERYTHROCYTE AND SUGAR SPECIFICITIES.
 DR PIR; PA0007; PA0007.
 KW Lectin; Glycoprotein.
 FT NON_TER 14 14
 SQ SEQUENCE 14 AA; 1732 MW; D804CE43B487C549 CRC64;

Query Match 21.0%; Score 21; DB 1; Length 14;
 Best Local Similarity 40.0%; Pred. No. 1.6e+03;
 Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 9 EQIEYFGDF 18
 DB 1 ETISFENQF 10

RESULT 12
 NUOS_SOLTU STANDARD; PRT; 23 AA.
 ID NUOS_SOLTU
 AC P80262;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE NADH-ubiquinone oxidoreductase 29 kDa subunit (EC 1.6.5.3)
 DE (EC 1.6.99.3) (Complex I-29KD) (CI-29KD) (Fragment).
 OS Solanum tuberosum (potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RN SEQUENCE.
 RP STRAIN=cv. Bintje; TISSUE=Tuber;
 RX MEDLINE=94124587; PubMed=8294484;
 RA Herz U., Schroeder W., Liddell A., Leaver C.J., Brennicke A.,
 Grohmann L.;
 RT "Purification of the NADH:ubiquinone oxidoreductase (complex I) of
 the respiratory chain from the inner mitochondrial membrane of
 Solanum tuberosum";
 RL J. Biol. Chem. 269:2263-2269(1994).
 CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
 CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
 TO BE UBIQUINONE.
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -1- CATALYTIC ACTIVITY: NADH + acceptor = NAD(+) + reduced acceptor.
 CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
 CC -1- SUBCELLULAR LOCATION: MATRIX SIDE OF THE MITOCHONDRIAL INNER
 MEMBRANE.
 DR PIR; C49732; C49732.
 KW Oxidoreductase; NAD; 23
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2653 MW; 4B41B8FF83412F58 CRC64;

```

Query Match      21.0%; Score 21; DB 1; Length 23;
Best Local Similarity 40.0%; Pred. No. 2.6e+03;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 12 EYFYG 16
    :|::|
Db 18 DYFYG 22

RESULT 13
UDP_LACCA
ID UDP_LACCA STANDARD; PRT; 23 AA.
AC P19662;
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Uridine phosphorylase (EC 2.4.2.3) (UDRPase) (Fragment).
GN UDP.
OS Lactobacillus casei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1582;
RN [1]
RP SEQUENCE.
RC STRAIN=ATCC 7469;
RX MEDLINE=90381286; PubMed=2119230;
RA Avraham Y., Grossowicz N., Yashpke J.;
RT "Purification and characterization of uridine and thymidine
RT phosphorylase from Lactobacillus casei.";
RL Biochim. Biophys. Acta 1040:287-293(1990).
CC -1- FUNCTION: THE ENZYMES WHICH CATALYZE THE REVERSIBLE PHOSPHORYLOSIS
CC OF PYRIMIDINE NUCLEOSIDES ARE INVOLVED IN THE DEGRADATION OF THESE
CC COMPOUNDS AND IN THEIR UTILIZATION AS CARBON AND ENERGY SOURCES,
CC OR IN THE RESCUE OF PYRIMIDINE BASES FOR NUCLEOTIDE SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: Uridine + phosphate -> uracil + alpha-D-ribose
CC 1-phosphate.
CC -1- PATHWAY: Nucleotide and deoxyribonucleotide catabolism.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE PNP/UDP FAMILY 1 OF PHOSPHORYLASES.
DR PIR: S11383; S11383.
DR InterPro: IPR000845; PNP_UDP.
DR POSITE; PS01232; PNP_UDP.1; PARTIAL.
KW Transferase; Glycosyltransferase.
FT UNSURE 1 1 OR S.
FT UNSURE 7 7 OR T.
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2630 MW; 918B2E2F32F35A17 CRC64;

Query Match      21.0%; Score 21; DB 1; Length 23;
Best Local Similarity 40.0%; Pred. No. 2.6e+03;
Matches 6; Conservative 2; Mismatches 3; Indels 4; Gaps 2;

QY 4 DTK-ICEQIEYFGD 17
    :|::|
Db 11 DVKLQCTDV---FGD 22

RESULT 14
PQQA_ACICA
ID PQQA_ACICA STANDARD; PRT; 24 AA.
AC P27532;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Coenzyme PQQ synthesis protein A (Coenzyme PQQ synthesis protein IV).
GN PQQA OR PQQIV.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LMD 79.41;

```

```

RX MEDLINE=89123056; PubMed=2536663;
RA Goosen N., Horsman H.P.A., Huinen R.G.M., van de Putte P.;
RT "Acinetobacter calcoaceticus genes involved in biosynthesis of the
RT coenzyme pyrrolo-quinoline-quinone: nucleotide sequence and
RT expression in Escherichia coli K-12.";
RL J. Bacteriol. 171:447-455(1989).
CC -1- FUNCTION: REQUIRED FOR COENZYME PYRROLO-QUINOLINE-QUINONE (PQQ)
CC BIOSYNTHESIS. THIS PROTEIN COULD FUNCTION AS A TEMPLATE FOR PQQ
CC SYNTHESIS AND PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH
CC ARE NECESSARY FOR THE SYNTHESIS.
CC -1- SIMILARITY: TO OTHER BACTERIAL PQQA.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X06452; -; NOT_ANNOTATED_CDS.
DR PIR: B32252; B32252.
KW PQQ.
SQ SEQUENCE 24 AA; 2938 MW; 0E671EB9B7EA6838 CRC64;

Query Match      21.0%; Score 21; DB 1; Length 24;
Best Local Similarity 33.3%; Pred. No. 2.8e+03;
Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 DTKICEQIEYF 15
    :|::|
Db 10 DLRIQFEVTMYF 21

RESULT 15
PQQA_PSEFL
ID PQQA_PSEFL STANDARD; PRT; 24 AA.
AC P55171;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Coenzyme PQQ synthesis protein A.
GN PQQA.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHAO;
RX MEDLINE=96064397; PubMed=8526497;
RA Schneider U., Keel C., Defago G., Haas D.;
RT "Tn5-directed cloning of pqq genes from Pseudomonas fluorescens CHAO:
RT mutational inactivation of the genes results in overproduction of the
RT antibiotic pyoluteorin.";
RL Appl. Environ. Microbiol. 61:3856-3864(1995).
CC -1- FUNCTION: REQUIRED FOR COENZYME PYRROLO-QUINOLINE-QUINONE (PQQ)
CC BIOSYNTHESIS. THIS PROTEIN COULD FUNCTION AS A TEMPLATE FOR PQQ
CC SYNTHESIS AND PROVIDE THE GLUTAMATE AND TYROSINE RESIDUES WHICH
CC ARE NECESSARY FOR THE SYNTHESIS.
CC -1- SIMILARITY: TO OTHER BACTERIAL PQQA.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X87299; CAA60732.1; -;
DR EMBL; X87299; CAA60731.1; ALT_INIT.
KW PQQ.

```

SQ SEQUENCE 24 AA; 2871 MW; ED13A879B2EA3E8E CRC64;

Query Match 21.0%; Score 21; DB 1; Length 24;
 Best Local Similarity 33.3%; Pred. No. 2.8e+03;
 Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 DTKICEQIEYYF 15

|:|::||

Db 10 DLRIGFEVTMYF 21

Search completed: April 23, 2003, 13:43:51
 Job time : 5.75169 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: April 23, 2003, 13:36:43 ; Search time 9.40449 Seconds
(without alignments)
183.999 Million cell updates/sec

Title: US-09-836-073-15
Perfect score: 100
Sequence: 1 LDLDTKICEQIEYVFGDF 18

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 4984

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	31	31.0	23	2 S13298	benzaldehyde dehyd
2	30	30.0	18	2 JS0647	pheromonotropin -
3	30	30.0	25	2 A60286	heat-stable serine
4	29	29.0	18	2 PH1368	Ig heavy chain DJ
5	29	29.0	25	2 C57001	endo-1,4-beta-xyla
6	29	29.0	25	2 PH1715	Ig heavy chain V r
7	27.5	27.5	23	2 PH1725	Ig heavy chain V r
8	27	27.0	12	2 PH0771	T-cell receptor be
9	27	27.0	14	2 PH1626	Ig H chain V-D-J r
10	27	27.0	14	2 PH1598	Ig H chain V-D-J r
11	27	27.0	18	2 PH1629	hemoglobin beta-2
12	27	27.0	19	2 A05305	Ig heavy chain DJ
13	27	27.0	22	2 PH1325	Ig heavy chain DJ
14	27	27.0	22	2 PH1359	Ig heavy chain DJ
15	27	27.0	23	2 S02201	prolamin alpha-3 -
16	27	27.0	24	2 PH1696	Ig heavy chain V r
17	26	26.0	16	2 E53284	T-cell receptor be
18	26	26.0	19	2 B53145	high conductance c
19	26	26.0	19	2 S75558	T cell receptor V-
20	25.5	25.5	25	2 PH1734	Ig heavy chain V r
21	25	25.0	15	2 PH1366	Ig heavy chain DJ
22	25	25.0	17	2 A51211	anantin - Streptom
23	25	25.0	18	1 MTFDFC	melanotropin beta
24	24.5	24.5	21	2 PH1730	Ig heavy chain V r
25	24.5	24.5	23	2 PH1707	Ig heavy chain V r
26	24.5	24.5	23	2 PH1722	Ig heavy chain V r
27	24.5	24.5	23	2 PH1724	Ig heavy chain V r
28	24.5	24.5	23	2 PH1727	Ig heavy chain V r
29	24.5	24.5	23	2 PH1723	Ig heavy chain V r

30	24.5	24.5	24	2 PH1710	Ig heavy chain V r
31	24.5	24.5	25	2 PH1716	Ig heavy chain V r
32	24.5	24.5	25	2 PH1733	Ig heavy chain V r
33	24	24	12	2 S57570	T cell receptor V-
34	24	24.0	14	2 S57569	T cell receptor V-
35	24	24.0	14	2 S57638	T cell receptor V-
36	24	24.0	16	1 MTFDFB	melanotropin beta
37	24	24.0	22	2 PH1678	Ig heavy chain V r
38	24	24.0	22	2 PH1679	Ig heavy chain V r
39	24	24.0	23	2 PH1681	Ig heavy chain V r
40	24	24.0	23	2 PH1682	Ig heavy chain V r
41	24	24.0	23	2 PH1694	Ig heavy chain V r
42	24	24.0	24	2 PH1683	Ig heavy chain V r
43	24	24.0	24	2 PH1685	Ig heavy chain V r
44	24	24.0	24	2 PH1698	Ig heavy chain V r
45	24	24.0	25	2 S29283	hydrogenase (EC 1.

ALIGNMENTS

RESULT 1

S13298

benzaldehyde dehydrogenase (EC 1.2.1.-) I - Acinetobacter calcoaceticus (fragment)
C/Species: Acinetobacter calcoaceticus
C/Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 17-Mar-1999
C/Accession: S13298

R/Chalmers, R.M.; Keen, J.N.; Fewson, C.A.

Biochem. J. 273, 99-107, 1991

A/Title: Comparison of benzyl alcohol dehydrogenases and benzaldehyde dehydrogenase

gene pathway in Pseudomonas putida.

A/Reference number: S13298; MUID:91113163; PMID:1989592

A/Accession: S13298

A/Molecule type: protein

A/Residues: 1-23 <CHA>

A/Experimental source: NCIB 8250

C/Keywords: oxidoreductase

Query Match 31.0%; Score 31; DB 2; Length 23;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 2; Gaps 1;

QY 2 LDLDTKICEQI--EYVF 15

DB 2 NIQTKIEQIWKHEIF 17

RESULT 2

JS0647

pheromonotropin - armyworm (Pseudaletia separata)

C/Species: Pseudaletia separata

C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 11-Jul-1997

C/Accession: JS0647

R/Matsumoto, S.; Fonagy, A.; Kurihara, M.; Uchiumi, K.; Nagamine, T.; Chifimatsu, D.

Biochem. Biophys. Res. Commun. 182, 534-539, 1992

A/Title: Isolation and primary structure of a novel pheromonotropic neuropeptide st

A/Reference number: JS0647; MUID:92134266; PMID:1734867

A/Accession: JS0647

A/Molecule type: protein

A/Residues: 1-18 <MAT>

C/Keywords: amidated carboxyl end

F18/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 30.0%; Score 30; DB 2; Length 18;

Best Local Similarity 38.5%; Pred. No. 2e+02;

Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEY 13

DB 2 LSYDDKVFENVVF 14

RESULT 3

A60286
heat-stable serine proteinase (EC 3.4.21.-) - Thermomonospora fusca (strain YX) (fragment)
N;Alternate names: YX-proteinase
C;Species: Thermomonospora fusca
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Dec-1996
C;Accession: A60286
R;Kristjansson, M.M.; Kinsella, J.E.
Int. J. Pept. Protein Res. 36, 201-207, 1990
A;Title: Heat stable proteinase from Thermomonospora fusca. Characterization as a serine
A;Reference number: A60286; MUID:91107200; PMID:2132918
A;Accession: A60286
A;Molecule type: protein
A;Residues: 1-25 <KRI>
C;Keywords: extracellular protein; hydrolase; serine proteinase

Query Match 30.0%; Score 30; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 13 YFGDF 18
|||||
Db 10 YFGNY 15

RESULT 4
PH1368
Ig heavy chain DJ region (clone C111-112) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PH1368
R;Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A;Title: Predominance of fetal type DJH joining in young children with B precursor lymph
A;Reference number: PH1302; MUID:93094761; PMID:1460419
A;Accession: PH1368
A;Molecule type: DNA
A;Residues: 1-18 <WAS>
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 29.0%; Score 29; DB 2; Length 18;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 IEYVFG 16
|||||
Db 5 MEYIYG 10

RESULT 5
C57001
endo-1,4-beta-xylanase (EC 3.2.1.8). 4 - Streptomyces roseisclerotici (fragment)
C;Species: Streptomyces roseisclerotici
C;Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 23-May-1997
C;Accession: C57001
R;Grabski, A.C.; Forrester, I.T.; Patel, R.; Jeffries, T.W.
Protein Expr. Purif. 4, 120-129, 1993
A;Title: Characterization and N-terminal amino acid sequences of beta-(1-4)endoxylanases
A;Reference number: A57001; MUID:93229899; PMID:8471845
A;Accession: C57001
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-25 <GRA>
A;Experimental source: strain NRRL B-11019
A;Note: sequence extracted from NCBI backbone (NCBIP:130009)
C;Function:
A;Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
A;Pathway: xylan degradation
C;Superfamily: Streptomyces endo-1,4-beta-xylanase A; Streptomyces endo-1,4-beta-xylanase
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation

Query Match 29.0%; Score 29; DB 2; Length 25;
Best Local Similarity 62.5%; Pred. No. 4.1e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 9 EQEYVFG 16
|||||
Db 10 QQSGYVFG 17

RESULT 6
PH1715
Ig heavy chain V region (clone ASC-13) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C;Accession: PH1715
R;McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A;Title: Antigen-driven B cell differentiation in vivo.
A;Reference number: PH1675; MUID:93301607; PMID:8315385
A;Accession: PH1715
A;Molecule type: mRNA
A;Residues: 1-25 <MCH>
A;Experimental source: B cell
A;Note: the authors translated the codon GTA for residue 11 as Thr and ACA for res
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 29.0%; Score 29; DB 2; Length 25;
Best Local Similarity 46.2%; Pred. No. 4.1e+02;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 DTKICEQIEYVFG 16
|||||
Db 7 NTKSVETRFYVYG 19

RESULT 7
PH1725
Ig heavy chain V region (clone GCC-5) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C;Accession: PH1725
R;McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A;Title: Antigen-driven B cell differentiation in vivo.
A;Reference number: PH1675; MUID:93301607; PMID:8315385
A;Accession: PH1725
A;Molecule type: mRNA
A;Residues: 1-23 <MCH>
A;Experimental source: B cell
A;Note: the authors translated the codon ACA for residue 13 as Ala
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 27.5%; Score 27.5; DB 2; Length 23;
Best Local Similarity 40.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 4 DTKICEQIEYVFGDF 18
|||||
Db 7 NTKSVE-TRYIYGSY 20

RESULT 8
PH0771
T-cell receptor beta chain (PES.1.1.1) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
C;Accession: PH0771
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility com
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0771

A:Molecule type: mRNA
 A:Residues: 1-12 <CAS>
 A:Cross-references: EMBL:X60865; NID:953624; PIDN:CAA43255.1; PID:g53625
 A:Experimental source: T lymphocyte
 C:Keywords: T-cell receptor

Query Match 27.0%; Score 27; DB 2; Length 12;
 Best Local Similarity 71.4%; Pred. No. 3.9e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 QIEYVFG 16
 | | | | |
 Db 6 QYEQYFG 12

RESULT 9

PH1626

Ig H chain V-D-J region (clone B-less 118) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999

C:Accession: PH1626
 R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
 J. Exp. Med. 178, 317-329, 1993

A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
 A:Reference number: PH1580; MUID:93301609; PMID:8315387
 A:Accession: PH1626
 A:Molecule type: DNA
 A:Residues: 1-14 <LEV>
 A:Experimental source: bone marrow pre-B lymphocyte
 C:Keywords: immunoglobulin

Query Match 27.0%; Score 27; DB 2; Length 14;
 Best Local Similarity 40.0%; Pred. No. 4.6e+02;
 Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 CEQIEYVFGD 17
 | : | | |
 Db 1 CMRAXXYGD 10

RESULT 10

PH1598

Ig H chain V-D-J region (wild-type clone 306) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999

C:Accession: PH1598
 R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
 J. Exp. Med. 178, 317-329, 1993

A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
 A:Reference number: PH1580; MUID:93301609; PMID:8315387
 A:Accession: PH1598
 A:Molecule type: DNA
 A:Residues: 1-14 <LEV>
 A:Experimental source: bone marrow pre-B lymphocyte
 C:Keywords: immunoglobulin

Query Match 27.0%; Score 27; DB 2; Length 14;
 Best Local Similarity 50.0%; Pred. No. 4.6e+02;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 8 CEQIEYVFGD 17
 | : | | |
 Db 1 CAKGYGGD 10

RESULT 11

PH1629

Ig H chain V-D-J region (clone B-less 155) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999

C:Accession: PH1629
 R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
 J. Exp. Med. 178, 317-329, 1993

A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
 A:Reference number: PH1580; MUID:93301609; PMID:8315387
 A:Accession: PH1629
 A:Molecule type: DNA
 A:Residues: 1-18 <LEV>
 A:Experimental source: bone marrow pre-B lymphocyte
 C:Keywords: immunoglobulin

Query Match 27.0%; Score 27; DB 2; Length 18;
 Best Local Similarity 36.4%; Pred. No. 6e+02;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 CEQIEYVFGDF 18
 | : | | |
 Db 1 CVRRDYDGSY 11

RESULT 12

A05305

hemoglobin beta-2 chain - Indian spiny-tailed lizard (fragment)
 C:Species: Uromastyx hardwickii (Indian spiny-tailed lizard)
 C>Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 01-Aug-1997

C:Accession: A05305
 R:Naqvi, S.; Zaidi, Z.H.; von Bahr-Lindstrom, H.; Carlquist, M.; Jornvall, H.
 FEBS Lett. 162, 290-295, 1983

A:Reference number: A91314; MUID:84029159; PMID:6628672
 A:Accession: A05305
 A:Molecule type: protein

A:Residues: 1-19 <NAQ>
 C:Superfamily: globin; globin homology
 C:Keywords: erythrocyte; oxygen carrier

Query Match 27.0%; Score 27; DB 2; Length 19;
 Best Local Similarity 80.0%; Pred. No. 6.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 14 YFGDF 18
 | : | | |
 Db 1 FFGDF 5

RESULT 13

PH1325

Ig heavy chain DJ region (clone C199-121) - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: PH1325
 R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
 J. Exp. Med. 176, 1577-1581, 1992

A:Title: Predominance of fetal type DJH joining in young children with B precursor-
 A:Reference number: PH1302; MUID:93094761; PMID:1460419
 A:Accession: PH1325
 A:Molecule type: DNA
 A:Residues: 1-22 <WAS>
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

Query Match 27.0%; Score 27; DB 2; Length 22;
 Best Local Similarity 66.7%; Pred. No. 7.4e+02;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 IEYVFG 16
 | : | | |
 Db 9 IHYVYG 14

RESULT 14

PH1359

Ig heavy chain DJ region (clone C178-121) - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C:Accession: PH1359
 R:Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.

J. Exp. Med. 176, 1577-1581, 1992
 A:Title: Predominance of fetal type DJH joining in young children with B precursor lymph
 A:Reference number: PH1302; MUID:93094761; PMID:1460419
 A:Accession: PH1359
 A:Molecule type: DNA
 A:Residues: 1-22 <WAS>
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin

Query Match 27.0%; Score 27; DB 2; Length 22;
 Best Local Similarity 44.4%; Pred. No. 7.4e+02;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 8 CEQIEYFG 16
 I : ||:
 Db 6 CYENYYYG 14

RESULT 15
 S02201
 prolamin alpha-3 - orchard grass (fragment)
 C:Species: Dactylis glomerata (orchard grass)
 C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 18-Jun-1993
 C:Accession: S02201
 R:Vvedenskaya, I.O.; Shlyapnikov, S.V.; Konarev, A.V.
 Biochemistry (N.Y.) 51, 1309-1311, 1986
 A:Title: Characterization of the N-terminal amino acid sequence of alpha-prolamine from
 A:Reference number: S02200
 A:Accession: S02201
 A:Molecule type: protein
 A:Residues: 1-23 <VVE>
 A:Note: this paper is a translation of the Russian paper published in Biokhimiya (1986)

Query Match 27.0%; Score 27; DB 2; Length 23;
 Best Local Similarity 42.9%; Pred. No. 7.8e+02;
 Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEYY 14
 : || ||:
 Db 2 VOLDPFFEQOQYY 15

Search completed: April 23, 2003, 13:48:54
 Job time : 9.40449 secs

RESULT 2
US-09-836-073-13
; Sequence 13, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836, 073

```
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Mouse
US-09-836-073-13
```

```
Query Match: 76.0%; Score 76; DB 9; Length 17;
Best Local Similarity 81.2%; Pred. No. 2.2e-05;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3 LDTKICEQIEYFGDF 18
   |: ||| ||||| |||||
Db 2 LEAKICHQIEYFGDF 17
```

RESULT 3

```
US-09-836-073-1
; Sequence 1, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-1
```

```
Query Match: 76.0%; Score 76; DB 9; Length 18;
Best Local Similarity 81.2%; Pred. No. 2.3e-05;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3 LDTKICEQIEYFGDF 18
   |: ||| ||||| |||||
Db 3 LEAKICHQIEYFGDF 18
```

RESULT 4

```
US-09-836-073-14
; Sequence 14, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bovine
US-09-836-073-14
```

```
Query Match: 76.0%; Score 76; DB 9; Length 18;
Best Local Similarity 81.2%; Pred. No. 2.3e-05;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 3 LDTKICEQIEYFGDF 18
   |: ||| ||||| |||||
Db 3 LEAKICHQIEYFGDF 18
```

RESULT 5

```
US-09-836-073-2
; Sequence 2, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-2
```

```
Query Match: 74.0%; Score 74; DB 9; Length 18;
Best Local Similarity 75.0%; Pred. No. 4.6e-05;
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 3 LDTKICEQIEYFGDF 18
   |: ||| ||||| |||||
Db 3 LEAQICQIEYFGDF 18
```

RESULT 6

```
US-09-836-073-4
; Sequence 4, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-4
```

```
Query Match: 72.0%; Score 72; DB 9; Length 18;
Best Local Similarity 80.0%; Pred. No. 9.3e-05;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 4 DTKICEQIEYFGDF 18
   |: ||| ||||| |||||
Db 4 EAKICHQIEYFGDF 18
```

RESULT 7

```
US-09-836-073-9
```

; Sequence 9, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; PRIOR FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO. 9
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-9

Query Match 70.0%; Score 70; DB 9; Length 18;
Best Local Similarity 80.0%; Pred. No. 0.00019;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGD 17
|: ||| ||||| |||
Db 3 LEAKICHQIEYFGD 17

RESULT 8

US-09-836-073-11
; Sequence 11, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-11

Query Match 68.0%; Score 68; DB 9; Length 18;
Best Local Similarity 75.0%; Pred. No. 0.00038;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGD 18
|: ||| ||||| |||
Db 3 LEAKICHQIEYFGD 18

RESULT 9

US-09-836-073-12
; Sequence 12, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-12

Query Match 68.0%; Score 68; DB 9; Length 18;
Best Local Similarity 75.0%; Pred. No. 0.00038;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGD 18
|: ||| ||||| |||
Db 3 LEAKICHQIEYFGD 18

RESULT 10

US-09-836-073-10
; Sequence 10, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-836-073-10

Query Match 67.0%; Score 67; DB 9; Length 18;
Best Local Similarity 75.0%; Pred. No. 0.00054;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGD 18
|: ||| ||||| |||
Db 3 LEAKICHQIEYFGD 18

RESULT 11

US-09-836-073-16
; Sequence 16, Application US/09836073
; Patent No. US20020173475A1
; GENERAL INFORMATION:
; APPLICANT: Dasgupta, Asim
; APPLICANT: Das, S.
; APPLICANT: Baidya, Narayan
; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION
; FILE REFERENCE: 220002054822
; CURRENT APPLICATION NUMBER: US/09/836,073
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 09/316,630
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Rat
US-09-836-073-16

Query Match 65.5%; Score 65.5; DB 9; Length 19;
Best Local Similarity 76.5%; Pred. No. 0.00097;

Matches 13; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 3 LDTKICEQI-EYFGDF 18
 : : ||| || |||||
 Db 3 LEAKICHOIEYYFGDF 19

RESULT 12

US-09-836-073-3
 ; Sequence 3, Application US/09836073
 ; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baldya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-3

Query Match 65.0%; Score 65; DB 9; Length 18;
 Best Local Similarity 68.8%; Pred. No. 0.0011;

Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 LDTKICEQI-EYFGDF 18
 : : ||| || |||||
 Db 3 LQAKICHOIQYFGDF 18

RESULT 13

US-09-836-073-5
 ; Sequence 5, Application US/09836073
 ; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baldya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-5

Query Match 64.0%; Score 64; DB 9; Length 18;
 Best Local Similarity 73.3%; Pred. No. 0.0016;

Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 DTKICEQI-EYFGDF 18
 : : ||| |||||
 Db 4 EQKICHOIEYYFGDF 18

RESULT 14

US-09-836-073-7
 ; Sequence 7, Application US/09836073
 ; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baldya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-7

Query Match 61.0%; Score 61; DB 9; Length 18;
 Best Local Similarity 73.3%; Pred. No. 0.0045;

Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LDTKICEQI-EYFGDF 17
 : : ||| |||||
 Db 3 LEAKICHOIEYQGD 17

RESULT 15

US-09-836-073-8

; Sequence 8, Application US/09836073

; Patent No. US20020173475A1

; GENERAL INFORMATION:

; APPLICANT: Dasgupta, Asim

; APPLICANT: Das, S.

; APPLICANT: Baldya, Narayan

; TITLE OF INVENTION: METHODS TO INHIBIT VIRAL REPLICATION

; FILE REFERENCE: 220002054822

; CURRENT APPLICATION NUMBER: US/09/836,073

; CURRENT FILING DATE: 2002-10-24

; PRIOR APPLICATION NUMBER: 09/316,630

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 18

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-836-073-8

Query Match 60.0%; Score 60; DB 9; Length 18;
 Best Local Similarity 68.8%; Pred. No. 0.0063;

Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 LDTKICEQI-EYFGDF 18
 : : ||| |||||
 Db 3 LEAKICHOIEYQFGDF 18

Search completed: April 23, 2003, 13:52:10
 Job time : 10.9101 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein : protein search, using sw model

Run on: April 23, 2003, 13:38:28 ; Search time 8.49438 seconds
(without alignments)
62.349 Million cell updates/sec

Title: US-09-836-073-15
Perfect score: 100
Sequence: 1 LDTKICEQIEYFGDF 18
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues
Total number of hits satisfying chosen parameters: 147762

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB	ID	Description
1	76	76.0	18	4	US-09-316-630-3		Sequence 3, Appli
2	76	76.0	18	4	US-09-316-630-4		Sequence 4, Appli
3	35	35.0	16	2	US-07-894-063A-22		Sequence 22, Appl
4	33	33.0	20	2	US-07-678-974D-17		Sequence 17, Appl
5	33	33.0	20	2	US-08-945-168-22		Sequence 22, Appl
6	33	33.0	25	1	US-08-238-163-6		Sequence 6, Appl
7	33	33.0	25	4	US-09-227-357-334		Sequence 334, Appl
8	32	32.0	22	1	US-07-928-930A-9		Sequence 9, Appli
9	32	32.0	22	1	US-08-288-568-9		Sequence 9, Appli
10	32	32.0	22	1	US-08-487-461-9		Sequence 9, Appli
11	32	32.0	22	1	US-08-432-691-9		Sequence 9, Appli
12	32	32.0	22	1	US-08-487-459-9		Sequence 9, Appli
13	31	31.0	19	1	US-07-920-597-9		Sequence 9, Appli
14	30	30.0	10	2	US-08-433-318A-140		Sequence 140, App
15	30	30.0	10	2	US-08-922-048-140		Sequence 140, App
16	30	30.0	10	4	US-09-111-681C-156		Sequence 156, App
17	30	30.0	10	5	PCT-US96-06270-140		Sequence 140, App
18	30	30.0	11	1	US-08-433-318A-124		Sequence 124, App
19	30	30.0	11	2	US-08-922-048-124		Sequence 124, App
20	30	30.0	11	4	US-09-111-681C-140		Sequence 140, App
21	30	30.0	11	5	PCT-US96-06270-124		Sequence 124, App
22	30	30.0	14	1	US-08-433-318A-108		Sequence 108, App
23	30	30.0	14	1	US-08-433-318A-171		Sequence 171, App
24	30	30.0	14	2	US-08-922-048-108		Sequence 108, App
25	30	30.0	14	2	US-08-922-048-171		Sequence 171, App
26	30	30.0	14	4	US-09-111-681C-124		Sequence 124, App
27	30	30.0	14	4	US-09-111-681C-187		Sequence 187, App

Sequence 108, App
Sequence 171, App
Sequence 92, Appl
Sequence 155, App
Sequence 92, Appl
Sequence 155, App
Sequence 108, App
Sequence 171, App
Sequence 155, App
Sequence 11, Appl
Sequence 7, Appl
Sequence 20, Appl
Sequence 7, Appl
Sequence 20, Appl
Sequence 47, Appl
Sequence 55, Appl

28 30 30.0 14 5 PCT-US96-06270-108
29 30 30.0 14 5 PCT-US96-06270-171
30 30 30.0 15 1 US-08-433-318A-92
31 30 30.0 15 1 US-08-433-318A-155
32 30 30.0 15 2 US-08-922-048-92
33 30 30.0 15 2 US-08-922-048-155
34 30 30.0 15 4 US-09-111-681C-108
35 30 30.0 15 4 US-09-111-681C-171
36 30 30.0 15 5 PCT-US96-06270-92
37 30 30.0 15 5 PCT-US96-06270-155
38 29 29.0 10 2 US-08-618-696-11
39 29 29.0 10 3 US-09-033-753-11
40 29 29.0 11 2 US-08-618-696-7
41 29 29.0 11 2 US-08-618-696-20
42 29 29.0 11 3 US-09-033-753-7
43 29 29.0 11 3 US-09-033-753-20
44 29 29.0 14 2 US-08-726-464B-47
45 29 29.0 14 4 US-09-298-924-55

ALIGNMENTS

RESULT 1
US-09-316-630-3
; Sequence 3, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316.630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: LAP
US-09-316-630-3

Query Match 76.0%; Score 76; DB 4; Length 18;
Best Local Similarity 81.2%; Pred. No. 4.9e-06;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYFGDF 18
|: ||| |||||
Db 3 LEAKICHQIEYFGDF 18

RESULT 2
US-09-316-630-4
; Sequence 4, Application US/09316630
; Patent No. 6291637
; GENERAL INFORMATION:
; APPLICANT: Das, Saumitra
; APPLICANT: Dasgupta, Asim
; TITLE OF INVENTION: INTERFERENCE WITH VIRAL IRES-MEDIATED TRANSLATION BY A
; TITLE OF INVENTION: SMALL YEAST RNA REVEALS CRITICAL RNA-PROTEIN INTERACTIONS
; FILE REFERENCE: 22000-20548.21
; CURRENT APPLICATION NUMBER: US/09/316.630
; CURRENT FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: 08/817,953

```
; PRIOR FILING DATE: 1997-10-06
; PRIOR APPLICATION NUMBER: 08/321,427
; PRIOR FILING DATE: 1994-10-11
; PRIOR APPLICATION NUMBER: 60/086,527
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: B-LAP
; OTHER INFORMATION: this peptide is biotinylated
US-09-316-630-4

Query Match 76.0%; Score 76; DB 4; Length 18;
Best Local Similarity 81.2%; Pred. No. 4.9e-06;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDTKICEQIEYEGDF 18
   |. ||| ||||| |||||
Db 3 LEAKICHQIEYEGDF 18

RESULT 3
US-07-894-063A-22
; Sequence 22, Application US/07894063A
; Patent No. 5980899
; GENERAL INFORMATION:
; APPLICANT: BERZORSKY, Jay A.
; APPLICANT: SHIRAI, Mutsunori
; APPLICANT: AKATSUKA, Toshitaka
; APPLICANT: FEINSTONE, Stephen M.
; TITLE OF INVENTION: PEPTIDE FOR STIMULATION OF CYTOTOXIC T
; TITLE OF INVENTION: LYMPHOCYTES SPECIFIC FOR HEPATITIS C VIRUS IN A MAMMAL
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/894,063A
; FILING DATE: 19920610
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/162/NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-894-063A-22

Query Match 35.0%; Score 35; DB 2; Length 16;
Best Local Similarity 38.5%; Pred. No. 16;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 LDTKICEQIEY 14
```

```
Db 2 DLGVRVCEKMALY 14
||.:||.:|

RESULT 4
US-07-678-974D-17
; Sequence 17, Application US/07678974D
; Patent No. 5629146
; GENERAL INFORMATION:
; APPLICANT: DILLNER, JOAKIM
; APPLICANT: DILLNER, LENA
; TITLE OF INVENTION: METHOD FOR DETECTION OF HUMAN PAPILLOMAVIRUS
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERMAN & AISENBERG
; STREET: 1730 RHODE ISLAND AVENUE, N.W.,
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-3186
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/678,974D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: AISENBERG, Irwin M.
; REGISTRATION NUMBER: 19,007
; REFERENCE/DOCKET NUMBER: SG19171
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-293-1404
; TELEFAX: 202-872-0493
; TELEX: 440 069 AIS UI
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-07-678-974D-17

Query Match 33.0%; Score 33; DB 1; Length 20;
Best Local Similarity 46.2%; Pred. No. 44;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LLDTKICEQIEY 13
||.:||.:|
Db 7 LDICTSICKYPDY 19

RESULT 5
US-08-945-168-22
; Sequence 22, Application US/08945168
; Patent No. 5989548
; GENERAL INFORMATION:
; APPLICANT: DILLNER, JOAKIM
; TITLE OF INVENTION: PEPTIDE-BASED VACCINE AGAINST PAPILLOMA
; TITLE OF INVENTION: VIRUS
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C. SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

Thu Apr 24 08:53:43 2003

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,168
FILING DATE: 18-DEC-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/SE96/00533
FILING DATE: 23-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9501512-9
FILING DATE: 24-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: ORLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 7752-0002-0 PCT
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-945-168-22

Query Match 33.0%; Score 33; DB 2; Length 20;
Best Local Similarity 46.2%; Pred. No. 44;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 LDLDTKICEQIEY 13
||: |||: ||
Db 7 LDICTSICKYPDY 19

RESULT 6
US-08-238-163-6
; Sequence 6, Application US/08238163
; Patent No. 5569830
; GENERAL INFORMATION:
; APPLICANT: BENNETT, Alan
; APPLICANT: LABAVITCH, John M.
; APPLICANT: POWELL, Ann
; APPLICANT: STOTZ, Henrik
; TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
; TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/238,163
; FILING DATE: 03-MAY-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 2307E-540
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-238-163-6
Query Match 33.0%; Score 33; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 56;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 4 DTKICEQIEYFFGD 17
||: ||: ||: |||
Db 7 DKVLLQIKKAFGD 20
RESULT 7
US-09-227-357-334
; Sequence 334, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; EARLIER FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950

;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,947
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,964
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/056,360
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,684
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,984
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/055,954
;; EARLIER FILING DATE: 1997-08-18
;; EARLIER APPLICATION NUMBER: 60/058,785
;; EARLIER FILING DATE: 1997-09-12
;; EARLIER APPLICATION NUMBER: 60/058,664
;; EARLIER FILING DATE: 1997-09-12
;; EARLIER APPLICATION NUMBER: 60/058,660
;; EARLIER FILING DATE: 1997-09-12
;; EARLIER APPLICATION NUMBER: 60/058,661
;; EARLIER FILING DATE: 1997-09-12
;; NUMBER OF SEQ ID NOS: 672
;; SOFTWARE: Patent in Ver. 2.0
;; SEQ ID NO 334
;; LENGTH: 25
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-227-357-334

Query Match 33.0%; Score 33; DB 4; Length 25;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 DLDTKICEQ 10
|||:|:|:
Db 1 DLDTKICEQ 9

RESULT 8
US-07-928-930A-9
; Sequence 9, Application US/07928930A
; Patent No. 5344822
; GENERAL INFORMATION:
; APPLICANT: Levine, Daniel M.; Parker, Thomas S.;
; APPLICANT: Rubin, Albert L.
; TITLE OF INVENTION: Methods Useful in Endotoxin
; TITLE OF INVENTION: Prophylaxis and Therapy
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,930A
; FILING DATE: 19920812
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5344822man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: ROGO 206
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:

;; LENGTH: 22 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
US-07-928-930A-9

Query Match 32.0%; Score 32; DB 1; Length 22;
Best Local Similarity 54.5%; Pred. No. 70;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 DLDTKICEQ 12
|||:|:|:
Db 5 DLDTKICEQ 15

RESULT 9
US-08-288-568-9
; Sequence 9, Application US/08288568
; Patent No. 5506218
; GENERAL INFORMATION:
; APPLICANT: Parker, Thomas S.; Levine, Daniel M.;
; APPLICANT: Rubin, Albert L.; Gordon, Bruce R.; Saal, Stuart D.
; TITLE OF INVENTION: Methods Useful in Endotoxin Based
; TITLE OF INVENTION: Prophylaxis and Therapy
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb
; MEDIUM TYPE: storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/288,568
; FILING DATE: 10-AUGUST-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07453
; FILING DATE: 9-AUGUST-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/928,930
; FILING DATE: 12-AUGUST-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5508218man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: ROGO 211
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-288-568-9

Query Match 32.0%; Score 32; DB 1; Length 22;
Best Local Similarity 54.5%; Pred. No. 70;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 DLDTKICEQ 12
|||:|:|:
Db 5 DLDTKICEQ 15

RESULT 10
US-08-487-461-9
; Sequence 9, Application US/08487461

us-09-836-073-15.lim25.ra1

Thu Apr 24 08:53:43 2003

```

; Patent No. 5587366
; GENERAL INFORMATION:
; APPLICANT: Parker, Thomas S.; Levine, Daniel M.;
; APPLICANT: Rubin, Albert L.; Gordon, Bruce R.; Saal, Stuart D.
; TITLE OF INVENTION: Methods Useful in Endotoxin Based
; TITLE OF INVENTION: Prophylaxis and Therapy
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb
; MEDIUM TYPE: storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,461
; FILING DATE: June 7, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07453
; FILING DATE: 9-AUGUST-1993
; APPLICATION DATA:
; APPLICATION NUMBER: 07/928,930
; FILING DATE: 12-AUGUST-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 567485man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: ROGO 211.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-487-461-9
; Query Match 32.0%; Score 32; DB 1; Length 22;
; Best Local Similarity 54.5%; Pred. No. 70;
; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 DLDTKICEQIE 12
DB 5 DLQTKLNENVE 15

; RESULT 11
; US-08-432-691-9
; Sequence 9, Application US/08432691
; Patent No. 5614507
; GENERAL INFORMATION:
; APPLICANT: Parker, Thomas S.; Levine, Daniel M.;
; APPLICANT: Rubin, Albert L.; Gordon, Bruce R.; Saal, Stuart D.
; TITLE OF INVENTION: Methods Useful in Endotoxin Based
; TITLE OF INVENTION: Prophylaxis and Therapy
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb
; MEDIUM TYPE: storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,459
; FILING DATE: 7-JUNE-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07453
; FILING DATE: 9-AUGUST-1993
; APPLICATION DATA:
; APPLICATION NUMBER: 07/928,930
; FILING DATE: 12-AUGUST-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 567485man D.

; RESULT 12
; US-08-487-459-9
; Sequence 9, Application US/08487459
; Patent No. 5674855
; GENERAL INFORMATION:
; APPLICANT: Levine, Daniel M.; Parker, Thomas S.;
; APPLICANT: Rubin, Albert L.; Gordon, Bruce R.; Saal, Stuart D.
; TITLE OF INVENTION: Methods Useful in Endotoxin Based
; TITLE OF INVENTION: Prophylaxis and Therapy
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb
; MEDIUM TYPE: storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,459
; FILING DATE: 7-JUNE-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07453
; FILING DATE: 9-AUGUST-1993
; APPLICATION DATA:
; APPLICATION NUMBER: 07/928,930
; FILING DATE: 12-AUGUST-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 567485man D.

```

REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: ROGO 211.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-487-459-9

Query Match 32.08; Score 32; DB 1; Length 22;
Best Local Similarity 54.58; Pred. No. 70;
Matches 6; Conservative 2; Mismatches 0; Indels 3; Gaps 0;

QY 2 DLDTKICQIE 12
|||:|:
Db 5 DLTQKLNENVE 15

RESULT 13
US-07-920-597-9
Sequence 9, Application US/07920597
Patent No. 5447915
GENERAL INFORMATION:
APPLICANT: Schreiber, Stuart
APPLICANT: Burakoff, Steven
TITLE OF INVENTION: Terminally-Blocked Antiviral Peptides
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Aligretti & Witcoff, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/920,597
FILING DATE: 19920828
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO pct/us91/01142
FILING DATE: 28-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: No. 5447915nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 91,174-E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-920-597-9

Query Match 31.08; Score 31; DB 1; Length 19;
Best Local Similarity 75.08; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 DLDTKICE 9
|||:
Db 5 DSDTYICE 12

RESULT 14
US-08-433-318A-140
Sequence 140, Application US/08433318A
Patent No. 5663144
GENERAL INFORMATION:
APPLICANT: Greene, Mark I.
TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND
METHODS OF USING THE SAME
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 5919764ris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows
SOFTWARE: WORDPERFECT 6
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/433,318A
FILING DATE: 03-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mark DeLuca
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: UPN-2106
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 140:
SEQUENCE CHARACTERISTICS:
LENGTH: 10
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-433-318A-140

Query Match 30.08; Score 30; DB 1; Length 10;
Best Local Similarity 71.4%; Pred. No. 61;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 CEQIEY 14
|||:|:
Db 2 CEQEWY 8

RESULT 15
US-08-922-048-140
Sequence 140, Application US/08922048
Patent No. 5919764
GENERAL INFORMATION:
APPLICANT: Greene, Mark I.
TITLE OF INVENTION: COMPOUNDS THAT BIND TO p185 AND
METHODS OF USING THE SAME
NUMBER OF SEQUENCES: 184
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 5919764ris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:

us-09-836-073-15.lim25.ra1

Thu Apr 24 08:53:43 2003

```

; MEDIUM TYPE: DISKETTE, 3.5 INCH
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WORDPERFECT 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/922,048
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/433,318
; FILING DATE: 03-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark Deluca
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: UPN-2106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 140:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-922-048-140

```

```

Query Match      30.0%; Score 30; DB 2; Length 10;
Best Local Similarity 71.4%; Pred. No. 61;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      8 CEQIEYV 14
      |||||
Db      2 CEQEWY 8

```

Search completed: April 23, 2003, 13:50:26
Job time : 8.49438 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 13:28:18 ; Search time 27.2022 Seconds
(without alignments)
88.173 Million cell updates/sec

Title: US-09-836-073-15

Perfect score: 100

Sequence: 1 LLDLTKICEQIEYFGDF 18

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 320064

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

```
1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	76	75.0	18	21	AA199200		Human la autoantigen
2	65	65.0	21	14	AA193394		La/SSB epitope 17.
3	36	36.0	18	20	AA193362		Human DPD protein
4	35	35.0	15	22	AA193385		Hepatitis C virus
5	35	35.0	15	22	AA194018		Hepatitis C virus
6	35	35.0	16	15	AA194523		HCV NS5 region (25
7	34	34.0	21	23	AA198576		Insulin/insulin-li
8	34	34.0	23	22	AA190924		VH ligand-binding
9	33	33.0	20	11	AA194974		Papilloma virus ty
10	33	33.0	20	23	AA197249		HPV16 L1 T-cell ep

11	32	32.0	22	15	AA198546	Sequence of amphi
12	31	31.0	8	14	AA193395	La/SSB epitope 24.
13	31	31.0	9	22	AA198777	HIV gp120 protein
14	31	31.0	9	22	AA198778	HIV gp120 protein
15	31	31.0	10	23	AA192202	Murine MC-1 antibo
16	31	31.0	12	22	AA198408	HIV gp120 protein
17	31	31.0	12	22	AA198823	HIV gp120 protein
18	31	31.0	12	22	AA198824	HIV gp120 protein
19	31	31.0	12	22	AA198825	HIV gp120 protein
20	31	31.0	12	22	AA198826	HIV gp120 protein
21	31	31.0	12	22	AA198827	HIV gp120 protein
22	31	31.0	15	22	AA198587	HIV gp120 protein
23	31	31.0	15	22	AA198588	HIV gp120 protein
24	31	31.0	15	22	AA198589	HIV gp120 protein
25	31	31.0	15	22	AA198590	HIV gp120 protein
26	31	31.0	15	22	AA198591	HIV gp120 protein
27	31	31.0	15	22	AA198592	HIV gp120 protein
28	31	31.0	15	22	AA198593	HIV gp120 protein
29	31	31.0	15	22	AA198594	HIV gp120 protein
30	31	31.0	18	22	AA198631	HIV gp120 protein
31	31	31.0	18	22	AA198632	HIV gp120 protein
32	31	31.0	18	22	AA198633	HIV gp120 protein
33	31	31.0	18	22	AA198634	HIV gp120 protein
34	31	31.0	18	22	AA198635	HIV gp120 protein
35	31	31.0	18	22	AA198636	HIV gp120 protein
36	31	31.0	18	22	AA198637	HIV gp120 protein
37	31	31.0	18	22	AA198638	HIV gp120 protein
38	31	31.0	18	22	AA198639	HIV gp120 protein
39	31	31.0	18	22	AA198640	HIV gp120 protein
40	31	31.0	18	22	AA198641	HIV gp120 protein
41	31	31.0	19	10	AA191234	CD4 anti-receptor
42	31	31.0	19	10	AA191233	CD4 anti-receptor
43	31	31.0	19	10	AA191290	Sequence of human
44	31	31.0	19	10	AA195334	Sequence of peptid
45	31	31.0	20	18	AA1942916	Immunogenic Hepati

ALIGNMENTS

RESULT 1
AA1952200
ID AA1952200 standard; peptide; 18 AA.

XX AA1952200;

DT 14-MAR-2000 (first entry)

XX Human la autoantigen peptide (LAP).

XX La autoantigen; LAP; internal ribosome entry site; IRES; translation;
KW viral replication; IRNA; antiviral agent; picornavirus; flavivirus;
KW coronavirus; hepatitis virus; rhabdovirus; adenovirus; coxsackie virus;
KW parainfluenza virus; poliovirus; rhinovirus; encephalomyocarditis virus;
KW foot-and-mouth disease virus; echo virus; infectious bronchitis virus;
KW vesicular stomatitis virus.

XX Homo sapiens.

XX WO9961613-A2.

XX 02-DEC-1999.

XX 21-MAY-1999; 99WO-US11281.

XX 22-MAY-1998; 98US-0086527.

XX (REGC) UNIV CALIFORNIA.

XX Das S, Dasgupta A;

XX WPI; 2000-062712/05.

PT New yeast inhibitory peptide useful for inhibiting viral protein
 PT translation and replication -
 XX
 PS Claim 5; Page 57; 81pp; English.
 CC This sequence is the La autoantigen binding domain (LAP). LAP is a
 CC cellular protein which binds to an internal ribosome entry site (IRES).
 CC The peptide is used to inhibit mRNA translation, viral protein
 CC translation or viral replication. Viral mRNA translation is initiated at
 CC the IRES and the LAP peptide prevents translation initiation factors
 CC from binding at the site. The peptide can be used alone or in combination
 CC with an inhibitor RNA (IRNA see A245200). The LAP peptide is useful as
 CC an antiviral agent, which works through the inhibition of mRNA
 CC translation, especially viral mRNA. Examples of viruses which can be
 CC inhibited are picornavirus, flavivirus, coronavirus, hepatitis A B or C
 CC viruses, rabdovirus, adenovirus, and parainfluenza virus, poliovirus,
 CC rhinovirus, coxsackie virus, encephalomyocarditis virus, foot-and-mouth
 CC disease virus, echo virus, hepatitis C virus, infectious bronchitis
 CC virus, duck and human hepatitis B virus, and vesicular stomatitis virus.
 CC The peptide also inhibits replication of the above viruses. The LAP
 CC peptide selectively inhibits viral protein translation, and is therefore
 CC not toxic to the host cell.

XX Sequence 18 AA;
 SQ
 Query Match 76.0%; Score 76; DB 21; Length 18;
 Best Local Similarity 81.2%; Pred. No. 3.8e-05;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDKICEQIEYFGDF 18
 !: ||| |||||
 Db 3 LEAKICHQIEYFGDF 18

RESULT 2
 AAR43394
 ID AAR43394 standard; peptide; 21 AA.
 AC AAR43394;

DT 12-MAY-1994 (first entry)

DE La/SSB epitope 17.

XX Linear; epitope; 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 KW nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 XX RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.

OS Homo sapiens.

XX WO9321223-A.

XX 28-OCT-1993.

XX 13-APR-1993; 93WO-US03484.

XX 13-APR-1992; 92US-0867819.

XX (OKLA) UNIV OKLAHOMA STATE.

XX Harley JB;

XX WPI; 1993-351658/44.

XX New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic
 PT lupus erythematosus

XX Claim 1; Page 30; 43pp; English.

XX The sequences given in AAR43391-562 are linear epitopes which are

CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B',
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polypyridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.

XX Sequence 21 AA;

Query Match 65.0%; Score 65; DB 14; Length 21;
 Best Local Similarity 91.7%; Pred. No. 0.0024;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ICEQIEYFGDF 18
 !: |||||
 Db 1 ICHQIEYFGDF 12

RESULT 3
 AAW93362
 ID AAW93362 standard; Protein; 18 AA.

AC AAW93362;

DT 28-MAY-1999 (first entry)

DE Human DPD protein fragment #1.

XX DPD; dihydropyrimidine-dehydrogenase; monoclonal antibody; MAB;
 KW immunoassay reagent; cancer patient; treatment; antitumor agent;
 KW 5-fluorouracil; affinity purification; toxicity.

XX Homo sapiens.

XX DE19837391-A1.

XX 25-FEB-1999.

XX 18-AUG-1998; 98DE-1037391.

XX 22-AUG-1997; 97EP-0114630.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Hasegawa M, Yoshikubo T;

XX WPI; 1999-155202/14.

XX Monoclonal antibody specific for dihydropyrimidine dehydrogenase -
 PT for assessing patient response to 5-fluorouracil antitumor agents

XX Disclosure; Page 24; 34pp; German.

XX This invention describes a monoclonal antibody (MAB) specific for
 CC dehydropyrimidine dehydrogenase (DPD). This MAB is used as immunoassay
 CC reagents to identify a lack of DPD in a patient and to assess the
 CC sensitivity of cancer patients to treatment with antitumor agents of the
 CC 5-fluorouracil (5-FU) type. The MAB can also be used for affinity
 CC purification of DPD. DPD is involved in reduction of 5-FU (and related
 CC catabolites and derivatives) and lack of it is associated with increased

CC toxicity of this type of antitumor agent. It has specific binding
CC interaction. The MAb provide a sensitive and reliable test for DPD,
CC which is simple, rapid and suitable for routine screening.
XX
SQ Sequence 18 AA;

Query Match 36.0%; Score 36; DB 20; Length 18;
Best Local Similarity 60.0%; Pred. No. 67;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 CEQIEYFGD 17
||:| |||
Db 1 CEKLENNFGD 10

RESULT 4
AAJ03385
ID AAJ03385 standard; Peptide; 15 AA.
XX
AC AAJ03385;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #3376.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US19774.
XX
PR 19-JUL-1999; 99US-0357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
PT A new composition useful as a vaccines against hepatitis C virus -
XX
PS Example 5; Page 177; 214pp; English.
XX

The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention.
XX
SQ Sequence 15 AA;

Query Match 35.0%; Score 35; DB 22; Length 15;
Best Local Similarity 38.5%; Pred. No. 79;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLDTKICQIEY 14
||:| ||:|
Db 3 DLGVRVCEKMA 15

RESULT 5
AAJ04018
ID AAJ04018 standard; Peptide; 15 AA.
XX
AC AAJ04018;
XX
DT 02-JUL-2001 (first entry)
XX

The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention.
XX
SQ Sequence 15 AA;

Query Match 35.0%; Score 35; DB 22; Length 15;
Best Local Similarity 38.5%; Pred. No. 79;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLDTKICQIEY 14
||:| ||:|
Db 3 DLGVRVCEKMA 15

RESULT 6
AAJ45253
ID AAR45253 standard; peptide; 16 AA.
XX
AC AAR45253;
XX
DT 29-JUN-1994 (first entry)
XX
DE HCV NS5 region (2583-2598).
XX
KW Cytotoxic T lymphocyte; hepatitis C virus; HCV; NS5 region;
KW amphipathicity; T-cell; epitope; stimulation; diagnosis;
KW prognosis; infection; vaccine;
XX
OS Synthetic (based on sequence of HCV1 Chiron isolate).
XX
PN WO9325575-A.
XX
PD 23-DEC-1993.
XX
PF 10-JUN-1993; 93WO-US05434.
XX
PR 10-JUN-1992; 92US-0894063.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Akatsuka T, Berzofsky JA, Feinstein SM, Shirai M;
XX
DR WPI; 1994-007453/01.

XX Hepatitis C virus epitope #4009.
DE
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US19774.
XX
PR 19-JUL-1999; 99US-0357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
PT A new composition useful as a vaccines against hepatitis C virus -
XX
PS Example 5; Page 198; 214pp; English.
XX

The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention.
XX
SQ Sequence 15 AA;

Query Match 35.0%; Score 35; DB 22; Length 15;
Best Local Similarity 38.5%; Pred. No. 79;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 DLDTKICQIEY 14
||:| ||:|
Db 3 DLGVRVCEKMA 15

RESULT 6
AAR45253
ID AAR45253 standard; peptide; 16 AA.
XX
AC AAR45253;
XX
DT 29-JUN-1994 (first entry)
XX
DE HCV NS5 region (2583-2598).
XX
KW Cytotoxic T lymphocyte; hepatitis C virus; HCV; NS5 region;
KW amphipathicity; T-cell; epitope; stimulation; diagnosis;
KW prognosis; infection; vaccine;
XX
OS Synthetic (based on sequence of HCV1 Chiron isolate).
XX
PN WO9325575-A.
XX
PD 23-DEC-1993.
XX
PF 10-JUN-1993; 93WO-US05434.
XX
PR 10-JUN-1992; 92US-0894063.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Akatsuka T, Berzofsky JA, Feinstein SM, Shirai M;
XX
DR WPI; 1994-007453/01.

XX Peptide for stimulation of cytotoxic T lymphocytes specific for
PT hepatitis C virus - used in diagnosis and prognosis of HCV
PT infection and in vaccines
XX
PS Disclosure; Fig 1A; 7lpp; English.
XX
CC Based on the sequence of the NS5 region of HCV, peptides (AAR44538-50
CC and AAR45245-61) were synthesised. The peptide sequences were
CC selected on the basis of amphipathicity as potential T-cell epitopes.
CC The peptides were then tested for ability to induce a cytotoxic T-
CC cell response against cells expressing HCV NS5 protein. A peptide
CC comprising a first segment of at least 8 consecutive residues from
CC the peptides given in AAR45246-48 is claimed.
CC The amphipathic score of the peptide below is 12.2.
XX
XX Sequence 16 AA;
SQ
Query Match 35.0%; Score 35; DB 15; Length 16;
Best Local Similarity 38.5%; Pred. NO. 85;
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 2 DLDTKICEQIEYY 14
|| :|||: |
DB 2 DLGVRVCEKVALY 14

RESULT 7
AAU88576
ID AAU88576 standard; Peptide; 21 AA.
XX
AC AAU88576;
XX
XX 18-JUN-2002 (first entry)
DE Insulin/insulin-like growth factor receptor-binding peptide #532.
XX
XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;
KW ophthalmological; insulin; receptor; gene therapy; diabetes;
KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;
KW diabetic retinopathy; neurological diseases; stroke;
KW diabetic neuropathy.
XX
XX Synthetic.
OS
XX WO200172771-A2.
PN
XX 04-OCT-2001.
PD
XX 29-MAR-2000; 2000WO-US08528.
PF
XX 29-MAR-2000; 2000WO-US08528.
PR
XX (DGIB-) DGI BIOTECHNOLOGIES LLC.
PA (NOVO) NOVO NORDISK AS.
XX
XX Beasley J, Blume AJ, Schaeffer L, Pillutla R, Brandt J;
PI Brissette R, Spetzel J, Cheng W, Ostergaard S, Mandeckl WS;
PI Hansen PH, Ravera M, Hsiao K;
XX
XX WPI; 2002-025774/03.
DR
XX
XX Modulating insulin activity in mammalian cells, for treating e.g.
PT diabetes and tumours, comprises using peptides that bind to insulin or
PT insulin-like growth factor receptors
XX
XX Disclosure; Figure 1E-1; 390pp; English.
PS
XX The invention relates to a method of modulating insulin activity in
CC mammalian cells by administering a peptide that binds the insulin
CC receptor (IR). A composition containing a peptide, optionally expressed
CC from gene therapy vectors, that binds to Site 1 of IR and an insulin
CC antagonist are useful for treating diabetes. Also, peptides that are

CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are
CC useful for treating insulin-like growth factor (IGF)-sensitive tumours
CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1
CC receptor agonists are useful for treating neurological diseases,
CC including stroke and diabetic neuropathy. The peptides are also useful in
CC screening for compounds that bind to IR or IGF-1 receptor, potential
CC therapeutics and research reagents. AAU88034-AAU90957 represent IR
CC and/or IGF-1 receptor-binding peptides and related amino acid sequences
CC of the invention.
XX
XX Sequence 21 AA;
SQ
Query Match 34.0%; Score 34; DB 23; Length 21;
Best Local Similarity 40.0%; Pred. NO. 1.6e+02;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 5 TKICEQIEYY 14
|| |::: |
DB 4 TKACQRVNFY 13

RESULT 8
AAU00924
ID AAU00924 standard; Peptide; 23 AA.
XX
AC AAU00924;
XX
XX 18-MAY-2001 (first entry)
DT
XX
DE VH ligand-binding domain variant CDR3 region #31.
XX
XX Human; monoclonal antibody A6; VH ligand-binding domain; variable region;
KW parental ligand binding molecule; PLBM; framework region; FR; CDR;
KW combinatorial library; diagnosis; therapy; surface protein; pesticide;
KW cell-receptor; carbohydrate; lipid; antibiotic; steroid; hormone; dye;
KW herbicide; polymer; immunoglobulin; CDR3 region; M2 immunoglobulin;
KW panning.
XX
XX Homo sapiens.
OS
XX WO200118058-A2.
PN
XX 15-MAR-2001.
PD
XX 07-SEP-2000; 2000WO-CA01027.
PF
XX 07-SEP-1999; 99CA-2282179.
PR
XX 04-NOV-1999; 99US-0163546.
PR
XX (NOVO-) NOVOPHARM BIOTECH INC.
PA
XX
XX Kaplan H, Entwistle J, Tanha J, Narang S, Dan M;
PI WPI; 2001-235191/24.
DR
XX
XX Combinatorial libraries including phage display library comprises
PT variants of immunoglobulin VH fragments which comprises the framework
PT regions of wild-type or modified immunoglobulin VH domain of human A6
PT antibody
XX
XX Disclosure; Page 25; 133pp; English.
PS
XX The sequence represents a variant CDR3 region of a parental VH
XX ligand-binding domain of human monoclonal antibody A6, isolated by
CC panning against M2 immunoglobulin. Variants of a parental ligand
CC binding molecule (PLBM) comprise an immunoglobulin (Ig) VH fragment
CC including the framework (FR) regions of the wild-type/modified Ig VH
CC domain of human monoclonal antibody A6. The variant differs from PLBM at
CC residues constituting part of one of the CDRs of PLBM. A library
CC expressing binding domains can be created by cloning a parental DNA
CC sequence that encodes a parental domain, in order to produce parental
CC clones, and replacing a variable region of the clones with a variant DNA
CC sequence. A library of genetic packages can then be generated, each

CC package having a surface and a surface protein binding domain expressed
CC by the variant DNA. Combinatorial libraries are useful for generating
CC forms of recombinant antibody fragments that specifically bind target
CC ligands including natural and synthetic molecules and macromolecules to
CC be used in diagnostic and/or therapeutic processes. Examples include
CC enzymes, cell-receptors, carbohydrates, lipids, and organic-based and
CC inorganic-based molecules such as antibiotics, steroids, hormones,
CC pesticides, herbicides, dyes and polymers.

XX Sequence 23 AA;

Query Match 34.0%; Score 34; DB 22; Length 23;
Best Local Similarity 45.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 CEQIEYVFGDF 18
| : : | : | : |
DB 11 CFWDYKFSDF 21

RESULT 9
AAR04974
ID AAR04974 standard; protein; 20 AA.

XX AC AAR04974;

XX 03-OCT-1990 (first entry)

XX Papilloma virus type 16 L1 peptide no. 17.

XX PV type 16; immunoglobulin A; L1; ELISA: cervical cancer.

XX Synthetic.

XX WO9004790-A.

XX 03-MAY-1990.

XX 30-OCT-1989; 89WO-SE00612.

XX 28-OCT-1989; 89SE-0003870.

XX (MEDS-) MEDSCAND AB.

XX Dillner J, Dillner L;

XX WPI; 1990-164122/21.

XX Detecting papilloma virus infections - by identifying specific
PT antibodies against partic. viral proteins or fragments, esp. for
PT rapid diagnosis of cervical cancer.

XX Claim 1; Page 30; 57pp; English.

XX The peptide is one of 66 overlapping peptides which together cover
CC the entire sequences of the L1 and L2 proteins of human PV type 16.
CC The peptide was found to be reactive with IgA Abs in the sera of
CC patients with type 16 cervical cancer. It can therefore be used in
CC immunoassays for the diagnosis of PV-associated neoplasia, esp.
CC carcinoma of the cervix uteri, and to assess the risk of development
CC of such a carcinoma.
CC See also AAR04958-R05023.

XX Sequence 20 AA;

Query Match 33.0%; Score 33; DB 11; Length 20;
Best Local Similarity 46.2%; Pred. No. 2.2e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEY 13
| : : | : | : |
DB 7 LDICTSICKYPDY 19

RESULT 10
ABP47249

ID ABP47249 standard; Peptide; 20 AA.

XX AC ABP47249;

XX 15-AUG-2002 (first entry)

XX HPV16 L1 T-cell epitope peptide P21 SEQ ID NO 21.

XX HPV; human papillomavirus; T-cell epitope; L1; E7; virucide; cytostatic;
KW vaccine; immune response.

XX Human Papillomavirus 16.

XX WO200244384-A2.

XX 06-JUN-2002.

XX 30-NOV-2001; 2001WO-EPI4037.

XX 01-DEC-2000; 2000DE-1059631.

XX (MEDI-) MEDIGENE AG.

XX Nieland J, Kaufmann A;

XX WPI; 2002-471833/50.

XX New T cell epitopes derived from human papilloma virus, useful for
PT detecting immune activation and in a vaccine against papilloma virus
XX Example 1; Page 35; 126pp; German.

XX The invention relates to T-cell epitopes (I) derived from human
CC papillomavirus (HPV). (I) have virucide and cytostatic activity and are
CC able to induce a cytotoxic T cell response or mediate T helper cell
CC function. (I) and also compounds (II) containing them, vectors that
CC express (I) or (II), cells that contain (I) or (II) and/or complexes
CC of (I) or (II) with another compound, are used to detect or induce an
CC immune response, for diagnosis or therapy, particularly in vaccines for
CC control of human papilloma virus infection. The present sequence is that
CC of a HPV T-cell epitope of the invention.

XX Sequence 20 AA;

Query Match 33.0%; Score 33; DB 23; Length 20;
Best Local Similarity 46.2%; Pred. No. 2.2e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 LDLDTKICEQIEY 13
| : : | : | : |
DB 2 LDICTSICKYPDY 14

RESULT 11
AAR48546

ID AAR48546 standard; peptide; 22 AA.

XX AC AAR48546;

XX 11-AUG-1994 (first entry)

XX Sequence of amphipathic peptide which forms a helical wheel,
DE designated APOA-IV consensus.

XX Andotoxin caused toxicity; therapy; peptide.

XX Synthetic.

XX WO9404177-A.

PD 03-MAR-1994.
 XX 09-AUG-1993; 93WO-US07453.
 PF 12-AUG-1992; 9205-0928930.
 PR (ROGO-) ROGOSIN INST.
 XX Levine DM, Parker TS, Rubin AL;
 PI WPI; 1994-082831/10.
 XX Treatment of endotoxin caused toxicity - by admin. of particle
 PT contg. peptide, which is not an apolipoprotein, and lipid which
 PT associates with the endotoxin.
 XX Claim 12; Fig 4I; 32pp; English.
 PS The inventors claim a method for treating a subject for endotoxin-
 XX caused toxicity which comprises administering a particle contg. a
 CC peptide which is not an apolipoprotein and a lipid. The endotoxin is
 CC an E.coli endotoxin or S. typhimurium endotoxin. The peptide is an
 CC amphipathic peptide which forms a helical wheel. The pref. peptide is
 CC designated 18A. Note, in the examples, the sequence of 18A has a
 CC Gly substd. for Glu at posn. 12. A method wherein the helical wheel-
 CC forming peptide is selected from the group of peptides in figure 4
 CC (AAR48538-46) is claimed.
 XX Sequence 22 AA;
 SQ Query Match 32.0%; Score 32; DB 15; Length 22;
 Best Local Similarity 54.5%; Pred. No. 3.5e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 OY 2 DLDTKICEQIE 12
 DB 5 DLDTKLNEVE 15
 RESULT 12
 AAR43395
 ID AAR43395 standard; peptide; 8 AA.
 AC AAR43395;
 XX 12-MAY-1994 (first entry)
 DT La/SSB epitope 24.
 DE Linear; epitope: 60 kD; Ro/SSA; La/SSB; autoantigen; E/F; G; 70 kD;
 XX nuclear ribonucleoprotein; nRNP; Sm B/B'; polypeptide; antigen; D;
 KW systemic lupus erythematosus; SLE; autoantibody; U4/U6; U5; B; B';
 KW RNA polymerase III; U1; U2; Sjogrens syndrome; SS; human; vaccine; ss.
 OS Homo sapiens.
 XX WO9321223-A.
 PN 28-OCT-1993.
 XX 13-APR-1993; 93WO-US03484.
 XX 13-APR-1992; 92US-0867819.
 PR (OKLA) UNIV OKLAHOMA STATE.
 PA Harley JB;
 PI WPI; 1993-351658/44.
 XX New linear epitope(s) for human auto-antibodies - from the
 PT Ro/SSA, La/SSB and Sm B/B' antigens and ribonucleoprotein, used
 PT for diagnosing and treating auto-immune disorders e.g. systemic

PT lupus erythematosus
 XX Claim 1; Page 30; 43pp; English.
 PS The sequences given in AAR43391-562 are linear epitopes which are
 CC derived from the 60 kD Ro/SSA peptide, the La/SSB autoantigen,
 CC the 70 kD nuclear ribonucleoprotein (nRNP) and the Sm B/B'
 CC polypeptide. These antigens are common in systemic lupus
 CC erythematosus (SLE) and closely related disorders. The Ro/SSA family
 CC of proteins has been shown to have several molecular forms which are
 CC defined by the molecular weight of the antigen identified. The major
 CC form has a molecular weight of 60 kD and two additional forms have
 CC molecular weights of 52 and 54 kD. La/SSB is also a member of this
 CC group of autoantibodies and binds small RNAs with a polypyridine
 CC terminus. La/SSB is bound by a third of the anti-Ro/SSA precipitin
 CC positive sera. La/SSB has been shown to be a 46-50 kD monomeric
 CC phosphoprotein which associates with RNA polymerase III transcripts.
 CC Anti-Sm antibodies precipitate snRNPs containing the U1, U2, U4/U6 and
 CC U5 RNA. Anti-Sm antibodies may be directed against one or a
 CC combination of the polypeptides: B (26 kD), B' (27 kD), D (13 kD),
 CC E/F (11 kD doublet) and G (less than 10 kD). These epitopes may be
 CC used for preventing, treating or screening autoimmune disorders,
 CC especially SLE or Sjogrens syndrome (SS). They bind to a human
 CC autoantibody and may therefore be used as vaccines.
 XX Sequence 8 AA;
 SQ Query Match 31.0%; Score 31; DB 14; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 14 YFGDF 18
 DB 1 YFGDF 5
 RESULT 13
 AAB89777
 ID AAB89777 standard; Peptide; 9 AA.
 AC AAB89777;
 XX 23-MAY-2001 (first entry)
 DT HIV gp120 protein binding peptide #870.
 DE Human chemokine receptor; CD4; HIV; glycoprotein 120; gp120; antagonist;
 KW replication; CCR5; CXCR4; CD4; STRL33.
 XX Synthetic.
 OS WO200116182-A2.
 PN 08-MAR-2001.
 XX 25-AUG-2000; 2000WO-US23505.
 XX 27-AUG-1999; 99US-0151270.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Saxinger C;
 PI WPI; 2001-244398/25.
 DR Novel polypeptides useful for treating HIV infection, have homology to
 XX regions of domains of human chemokine receptors CCR5, CXCR4 and STRL33,
 PT and binds to HIV gp120 under physiological conditions
 XX Example 10; Page 67; 114pp; English.
 PS The present invention describes a number of peptides which are able to
 CC bind to HIV glycoprotein 120 (gp120). These are similar to the human